

***City Slicker* and Other Stylistically Restricted Words —**

A Corpus-Assisted Study of Stylistic Variation in Schoolbooks Used for  
Teaching English in Finnish Secondary Schools

Sanna Ax

University of Tampere

School of Language, Translation and Literary Studies

Master's Programme in English Language and Literature

Pro Gradu Thesis

October 2015

Tampereen yliopisto  
Kieli-, käännös- ja kirjallisuustieteiden yksikkö  
Englannin kielen ja kirjallisuuden maisteriopinnot

AX, SANNA: *City Slicker* and Other Stylistically Restricted Words — A Corpus-Assisted Study of Stylistic Variation in Schoolbooks Used for Teaching English in Finnish Secondary Schools

Pro gradu -tutkielma, 69 sivua + liitteet 4 kpl (35 sivua)  
Lokakuu 2015

Tässä pro gradu -tutkielmassa tarkastellaan tyylillisen vaihtelun ymmärtämisen merkitystä kielitaidolle ja selvitetään, kuinka tyylillinen vaihtelu ja sanaston kerrostuneisuus otetaan huomioon englannin opetuksessa Suomessa. Opetuksen sisällön tutkimiseen käytettiin oppikirjoja, joiden kappalesanastojen sisältämiä tyylillisesti rajoittuneita sanoja käytettiin indikaattoreina kirjojen sisältämästä tyylillisestä vaihtelusta. Tutkimuksen kolmen keskeisimmän tutkimuskysymyksen avulla selvitettiin sisältävätkö oppikirjat tyylillisesti rajoittuneita sanoja, käykö kirjoista ilmi sanojen käyttöön liittyvät rajoitukset ja kuinka yleisesti kyseisiä sanoja käytetään autenttisessa kielessä.

Teoriaosuudessa käsitellään vieraan kielen oppimiselle asetettuja kansallisia tavoitteita, kielen tyylillistä vaihtelua ja esitellään tyylin tasot *formaali*, *informaali* ja *slangi*. Lisäksi perehdytään sanaston oppimiseen vaikuttaviin tekijöihin ja korpuslingvistiikkaan sekä sen käyttömahdollisuuksiin kieltenopetuksessa.

Tutkimuksen ensimmäisen vaiheen aineistona käytettiin kolmea Suomessa yleisesti käytössä olevaa yläkoulun oppikirjasarjaa. Aineiston laajuutta rajattiin valitsemalla jokaisesta kirjasarjasta kaksi kappalekohtaista sanastoa jokaisen luokka-asteen teksti- tai tehtäväkirjasta. Kappaleiden valinta pohjautui vuoden 2004 perusopetuksen opetussuunnitelman perusteissa kirjattuihin englanninopetuksen teemoihin. Kirjojen kappalesanastoista etsittiin sanakirjoissa formaaleiksi, informaaleiksi tai slangiksi luokiteltuja sanoja.

Tutkimuksen toisen vaiheen pääasiallisena aineistona käytettiin kattavaa amerikanenglanninkielistä korpusta oppikirjoista löydettyjen sanojen käytön yleisyyden arviointiin ja sanoille ominaisten kontekstien määrittämiseen. Lisäksi käytettiin internetkieleen pohjautuvaa korpusta selvittämään amerikanenglannissa hyvin harvoin esiintyneiden sanojen yleisyyttä muissa maailmanenglanneissa.

Oppikirja-analyysin perusteella voidaan todeta, että tyylillisesti rajoittuneiden sanojen osuus oppikirjoissa on pieni, sanojen käännökset eivät useinkaan ilmaise sanan tyylillisiä rajoituksia, eikä tyyleiltään erilaisia sanoja pääsääntöisesti merkitä tai eroteta toisistaan kappalesanastoissa. Suurin osa tutkittujen oppikirjojen sanoista havaittiin neutraaleiksi eikä niitä voitu jakaa kategorioihin *formaali*, *informaali* tai *slangi*. Löydettyistä tyylillisesti rajoittuneista sanoista puolestaan suurin osa oli informaaleja ja oppikirjojen huomattiin eroavan toisistaan huomattavasti niiden sisältämien tyylillisesti rajoittuneiden sanojen määrässä.

Korpustutkimus paljasti useimpien oppikirjoista löytyneiden tyylillisesti rajoittuneiden sanojen olevan hyvin harvinaisia luonnollisessa kielenkäytössä ja tältä osin englannin oppikirjojen autenttisuus joutui kyseenalaiseksi. Niin tämän pro gradu -tutkielman kuin aiempienkin tutkimusten tulokset kehottavat uudelleenarvioimaan harvinaisten sanojen opettamisen hyödyt peruskoulussa ja korostavat sanojen käyttöön liittyvien rajoitusten opettamisen tärkeyttä.

Avainsanat: kielten opetus, tyyli, sanasto, oppikirjat, korpuslingvistiikka

# Table of Contents

<b>1. Introduction.....</b>	<b>1</b>
<b>2. Background.....</b>	<b>4</b>
2.1 National objectives for foreign language learning .....	4
2.2 Style, formality and vocabulary learning .....	5
2.3 Stylistic labels in dictionaries .....	9
2.3.1 Formal words and the label <i>formal</i> .....	10
2.3.2 Informal words and the label <i>informal</i> .....	11
2.3.3 Slang words and the label <i>slang</i> .....	11
2.4 Corpus linguistics, frequency and language teaching .....	13
<b>3. Material and methods.....</b>	<b>17</b>
3.1 Data collection and processing: Schoolbooks .....	17
3.1.1 The three book series .....	17
3.1.2 Themes and chapters.....	18
3.1.3 Learner's dictionaries .....	20
3.1.4 Analysing word lists .....	21
3.2 Data collection and processing: Corpora.....	22
3.2.1 COCA – <i>Corpus of Contemporary American English</i> .....	23
3.2.2 GloWbE – <i>Corpus of Web-Based Global English</i> .....	28
3.2.3 Frequency calculations and organising the data .....	28
<b>4. Results and analysis .....</b>	<b>31</b>
4.1 Stylistically restricted words in schoolbooks .....	31
4.1.1 Stylistically restricted words: 7 <sup>th</sup> grade .....	31
4.1.2 Stylistically restricted words: 8 <sup>th</sup> grade .....	34
4.1.3 Stylistically restricted words: 9 <sup>th</sup> grade .....	38
4.1.4 Summary of findings .....	41
4.2 Formal, informal and slang words in corpora .....	46
4.2.1 Frequencies of the words in COCA.....	46

4.2.2 Distribution of the frequent words in sub-corpora .....	49
4.2.3 Distribution of frequent and infrequent words in schoolbooks .....	51
4.2.4 Infrequent words in GloWbE.....	53
4.2.5 Summary of findings .....	56
<b>5. Discussion.....</b>	<b>58</b>
5.1 Answering the research questions .....	58
5.1.1 Do schoolbooks contain stylistically restricted words? .....	58
5.1.2 Does their restricted nature become clear to the user of the book? .....	59
5.1.3 How widely are the words used in authentic language? .....	60
5.2 Limitations and implications .....	61
5.2.1 Limitations of the study .....	61
5.2.2 Implications of the results and tasks for the future .....	62
<b>6. Conclusion.....</b>	<b>64</b>
<b>References .....</b>	<b>66</b>
<b>Appendices</b>	

## 1. Introduction

Finland came fourth in a recent ranking of 63 countries by adult English skills (EF Education First 2014a and Yle Uutiset 2015) and the proficiency of English as a foreign language can with confidence be said to be very high among Finns. The ranking is based on the EF English Proficiency Index which is calculated using data from tests containing grammar, vocabulary, reading and listening exercises (EF Education First, 2014b) which are the areas traditionally focused on in foreign language teaching. Even though tests of this kind can be very ingeniously devised and bring forth valuable information about language skills, I believe that some areas where improvement is actually needed become invisible if rankings and test results are accepted as absolute truths about language proficiency. If the rankings are too highly valued, the pressure to succeed in tests might hide the need of focusing on the aspects of language which are difficult to test. This might result in poorer teaching contents if what is tested is let to define what is being taught.

In my experience as an English teacher, although learners may perform well in language tests, reading student essays and hearing Finnish people speak English in various contexts has showed me that finding the appropriate level of formality in a foreign language is often problematic; overly complicated vocabulary and sentence structures are used in casual conversation and *gonnas* and *wannas* occur repeatedly in literature reviews by students even in the last grade of secondary school. As I argue later in the thesis, using incorrect or mismatching styles or failing to recognise the level of formality others use, can hinder communication and create negative impressions of the speaker or writer. The central role of register variation has been recognised in a variety of linguistic disciplines. In Systemic Functional Linguistics, for example, as Martin (2013, 23-24, 33-34) claims, different types of language, or *genres*, are used to encode knowledge in school as well as outside the classroom, and if students are unfamiliar with the type of language used, knowledge will not be acquired. Hence, I agree with Andersson and Trudgill (1992, 171) in that “[d]iscussing the use of different styles of speaking and writing for different purposes should form a central part of English-teaching, particularly in secondary schools.” Struggling with stylistic issues appears to be common among

learners and yet there is relatively little knowledge available on how to improve teaching methods and teaching materials in order to answer to the need.

The topic of this thesis is to discover how different styles are represented in language teaching materials. I chose schoolbooks used for teaching English as a foreign language in Finnish secondary schools to demonstrate the situation. The decision to examine chapter-specific word lists (or glossaries) was made to increase the objectivity of the study by producing quantitative, comparable data instead of qualitative analyses of texts and sentences. Secondary school books were chosen over primary school books because they comprise more advanced topics and texts, and are thus more likely to contain examples of different styles. Their word lists suit the purposes of the study well, too, for they introduce learners to a good amount of new vocabulary but still contain some basic words, too. I searched the word lists for stylistically marked words which are here interpreted to be indicators of stylistic variation in the books. My aim was to find out whether the books contain words of this kind and whether they guide learners of English to use these words correctly and in proper contexts. The main focus was on words which can be classified as *formal*, *informal* or *slang*, for the reason that these three levels of formality cover most of the contexts needed in the globalising world; letters, applications, reports, e-mails, text messages, social media, video conferences et cetera.

Realising the difficulties foreign language learners have in adapting their language to the context raised the question as to whether stylistic differences receive the attention they deserve in English teaching in Finland. My research questions rests on the idea that the significant role of schoolbooks in language classes allows that schoolbook analysis is used for gathering information on language teaching. The three research questions answered in this thesis are:

1. Do schoolbooks contain stylistically restricted words?
2. Does their restricted nature become clear to the user of the book?
3. How widely are the words used in authentic language?

As is apparent from the last two research questions, my hypothesis is that words of different levels of formality are represented in the books, but to what extent and what type of words is also under

examination. The contexts in which a word can or cannot occur are essential to the word's meaning, hence the second research question on these restrictions of use. The third question, on the frequency of the words, was composed for analysing the usefulness of the words studied and for estimating how authentic the language in the schoolbooks is.

To answer the questions above, this study combines methods from textbook studies, lexicography and corpus studies: First, schoolbooks are examined for the types of words they contain in order to find stylistically restricted words. Second, how the words are labelled in learner's dictionaries is used as basis for categorising the words by their level of formality and as a tool for analysing how successfully the words are translated in the schoolbooks. Finally, the usefulness of the words found and the authenticity of the language in the schoolbooks are evaluated against the frequency information gathered from corpora. I believe that the results of this study contribute new ideas to pedagogical corpus applications, language teaching and materials development and draw attention to the importance of stylistic differences in foreign language learning.

## 2. Background

In this section, I explain how the content of language teaching in Finland is regulated on the national level and how stylistic differences are taken into consideration in the national curriculum for foreign language teaching. I will also argue for the importance of learning about the different levels of formality and show how the issue has been addressed in literature. I will discuss style and vocabulary learning from the viewpoints of schoolbook writers, teachers, and learners before taking a deeper look at the definitions of the three labels, namely formal, informal, and slang, used in this study to categorise words in the glossaries of schoolbooks. Since dictionaries are consulted in the present study for more information on the formality of the words found in schoolbooks, I find it important to illuminate some aspects of the process of compiling dictionaries and assigning stylistic labels to words in connection with the discussion of the three stylistic labels of *formal*, *informal* and *slang*. How dictionaries differ from each other in the labelling of words is discussed briefly, too. The last subsections introduce corpora and corpus linguistics as a tool for analysing and developing language teaching materials.

### 2.1 National objectives for foreign language learning

Investing in education is seen as a way to maintain and further improve the competitiveness of the Finnish society and the wellbeing of its citizens (Finnish National Board of Education, 2014b). Free education is therefore provided for everyone and basic education is government-funded in Finland (Finlex 1998, 20). To guarantee equal opportunities for everyone, basic education is governed by a national core curriculum on the basis of which the local providers of basic education compose their own curricula (Finnish National Board of Education, 2014a). The National Core Curriculum for Basic Education (hereafter NCCBE) is determined by the Finnish National Board of Education (hereafter FNBE) which is a subordinate to the Ministry of Education and Culture (Finnish National Board of Education, 2014a and 2014c). Among several other matters related to the quality and quantity of



education, the NCCBE includes the objectives and core contents of the subjects taught during the nine years of basic education (Finlex 1998, 4-6 and Finnish National Board of Education, 2014a).

The objectives set in the NCCBE for English language proficiency in secondary school, i.e. grades seven, eight, and nine, include that pupils ought to become aware of the differences between the different variants of English and be prepared for slightly more demanding informal communication (Opetushallitus 2004, 141; my translation). Moreover, pupils should learn to communicate in various everyday situations as is considered acceptable in the English speaking countries (ibid.) Since socially acceptable communication is one of the nationally set objectives for foreign language learning in Finnish secondary schools, and the linguistic variation mentioned is not restricted to the geographical variants, it seems apt to expect that words representing different social and stylistic variants can be found in schoolbooks used for teaching English.

## 2.2 Style, formality and vocabulary learning

The term *style* can sometimes be used in a fairly broad sense and its connotations might differ between people. In this study, *style* is not used to cover the concepts of register, accent or social or regional dialects. As Andersson and Trudgill (1992, 172) point out, there is no logical or necessary dependency between these aspects of language variation, since “[i]t is possible to speak Standard English with Liverpool accent[ and i]t is possible to swear and use slang while speaking Standard English. All dialects are capable of being spoken in a whole range of formal and informal styles”. Therefore, the term *style* is used here to refer only to the variation along the scale of formality and, similarly to Andersson and Trudgill, *slang* is only used to refer to “words or uses of words or expressions which are extremely informal” (ibid., p. 16). Words whose use is clearly restricted to formal or informal language or slang are hence referred to as *stylistically restricted words* in this study. To make a clear distinction between two concepts, words specific to a certain regional variety are referred to as *regionally restricted words*.

The concept of style is important every time a way of writing or speaking is being discussed. The fact cannot be dismissed that “[w]hen we speak and write, we adjust our style to the context and audience of communication” (Jackson and Zé Amvela 2000, 183). If we fail to do this or fail to recognise the style others are using, the consequences can vary from sounding ridiculous to offending the reader or listener. As Andersson and Trudgill (1992, 191) state, language users “will always be judged both by what [they] say and how [they] say it.” Therefore, the importance of teaching learners about stylistic differences cannot be stressed enough.

There are several factors which can have an effect on style, but according to Andersson and Trudgill (*ibid.*, p. 171) in English “[f]ormality versus informality is most often signalled [...] through vocabulary”. Jackson and Zé Amvela (2000, 138) agree by stating that the most notable

[p]art of the distinctiveness of a style is achieved by the choice of vocabulary. A more formal context requires ‘formal’ vocabulary; an informal context will allow ‘colloquial’ vocabulary, perhaps ‘slang’; a very informal context may even allow the use of ‘taboo’ vocabulary.

Schoolbooks include various types of texts, for example letters, written conversations, articles and short stories, which follows the view that “[i]deally materials at all levels should provide frequent exposure to authentic input which is rich and varied” (Tomlinson 1998, 13). The themes and the contexts of these texts vary; hence the style of the vocabulary should vary accordingly for the material to appear authentic. Several authors agree that the language used for pedagogical purposes cannot (and should not) be authentic in the sense of being unedited and occurring in its natural (cultural and situational) context, but for the learning to be motivational the language of schoolbooks ought to be structured authentically on the model of the natural language use found in corpora (see Carter 1998b, 52-53, Mishan 2004, 11-15, 25-26, and Tomlinson 1998, 13). Hence, the texts can be edited or constructed for the purposes of the learners, in other words, modified to correspond to the target level of competence but as Tomlinson (*ibid.*) argues, “the input should vary in style, mode, medium and purpose and should be rich in features which are characteristic of authentic discourse in target language”. He continues that for the learners “to be able to use the language for general communication” they should be “exposed to planned, semi-planned and unplanned discourse” (*ibid.*)

i.e. to the different levels of formality. For this reason, examining the vocabulary taught and the possible style labels can provide us with a deeper understanding of the level of formality of the texts in the schoolbooks and of the ways of teaching sensitivity to stylistic variation.

The difficulties learners encounter with vocabulary are often associated with connotations and the levels of formality. In other words, learners who have mostly familiarized themselves with written language and fairly formal texts tend to use overly formal vocabulary or blend the different styles. Based on the written and spoken language used by students, Fox (1998, 27) points out that the familiar argument that common words need less teaching is not valid. Fox claims that the frequency of exposure to these words is not enough and that the language of students is “stilted, too formal and too high-level” probably due to “the lack of attention that has been drawn to them in the classroom” (ibid.). Also, compared to native speakers of English, Fox has observed that students do not use the most common words as frequently nor in as many contexts (ibid.). Based on these remarks, my hypothesis is that less common words are taught in language classes, hence the word lists in the schoolbooks used most probably contain words which are not neutral in style and can therefore be restricted in their uses.

Language teaching is constantly evolving, which requires that teachers and textbook writers modify their practices accordingly. In the 1970’s, Brumfit et al. (1978, 196) suggested that the teacher should “develop an awareness of different styles of English” and help in developing “sensitivity to appropriate use in different social situations”. To do this, teachers were told to encourage students “to assign style labels to [words]” (ibid.) as they first appear in traditional textbooks and wordlists. However, the materials used in teaching have come a long way since these instructions were given. As Littlejohn (1998, 190) notes, schoolbooks are not merely texts with questions and translation exercises anymore, but rather “complete ‘packages’ [...] with precise indications of the work that teachers and students are to do together”. These ‘packages’ often include a textbook and an exercise book for the students; a teacher’s file with answers, tests, extra exercises and suggestions on how to approach each theme and chapter; CDs with audio versions of the texts together with listening

exercises; and even access to online exercises. In this study the term *schoolbook* is used for the students' material which often contains a textbook and an exercise book and the term *word list* refers to the chapter specific glossaries in the books.

Given the fact that the materials are compiled to contain nearly everything students are expected to learn and the amount of money invested in these materials by schools, it is understandable that teachers make use of the materials as much as possible and that there is a significant increase in "[t]he extent to which materials now effectively structure classroom time" (ibid.). Since the schoolbooks have such an immense role in what is being taught and how, some of the responsibility of teaching "awareness of different styles of English" assigned to teachers by Brumfit et al. in 1978 (p. 196) might now be expected to lie on the shoulders of the authors of schoolbooks. It would therefore ease the task of the teacher if stylistically restricted words were somehow marked in the schoolbooks. Students could pay attention to the words when studying at home and ask clarifying questions during the lessons.

The fact that students are often left unaware of the different connotations a word might have, may mainly be due to the limited space in schoolbooks and the limited nature of the traditional word lists in general. As Carter (1998a, 29) states, "[w]ord lists can be very valuable, but however ingeniously devised, they are no more than lists of words and should be treated with caution." Although he (ibid., pp. 212–213) admits word lists to be necessary at the beginning of the language studies but as for the more advanced learners, Carter emphasises the advantages gained when the vocabulary is learned in context. Since it is impossible to know the requisite level of proficiency, he recommends that vocabulary is taught both with word lists and in context. According to Carter's (ibid., p. 214) view, words are not learned separately but in groups and the original context supports vocabulary learning by offering natural collocates, as opposed to word lists. Of course, for the *natural* collocates to occur, the texts in schoolbooks should be authentic. Moilanen (2002, 202), too, stresses the importance of seeing and using the words in contexts other than the texts in schoolbooks since this improves memorising and learning them. He continues that it is motivational for the learner if

the value of the word in terms of different contexts is made explicit: is the word needed in informal conversations or does the word mainly occur in writing (ibid.)

### 2.3 Stylistic labels in dictionaries

Computers and digitalisation have had a great impact on research methods on several fields of science and this is true to the study of languages, too. As Fox (1998, 25) states, the use of computers has made it possible to compile ever larger corpora for researchers to use in analysing, for example, how language is or was used in specific time, place or context. This has also changed the way of compiling dictionaries; intuition has been replaced by the analysed data of how language is actually used (ibid.). Therefore, the information the modern dictionaries can provide us with about words and their different uses and senses is often very accurate.

Dictionaries use labels to inform the reader if the use of the word is restricted to certain groups of people, geographical areas, or social contexts. As Jackson and Zé Amvela (2000, 138) note, most of the words in dictionaries are not labelled since they are considered “to be ‘neutral’ in their formality” and the words which are labelled tend to be “towards the extremes of the styles”. It is known that there are more terms labelled *informal* than *formal* in dictionaries (Jackson and Zé Amvela 2000, 138 and Jackson 2002, 111) and it is therefore of interest to find out how similar (or dissimilar) schoolbooks and their word lists are to dictionaries in this respect.

There is considerable variation between dictionaries in the division of the labels into categories and in the names given to both the categories and the labels. To give a few examples, *Macmillan English Dictionary for Advanced Learners* (MEDAL) divides the labels into ‘Style and attitude labels’ (e.g. *formal* and *showing approval*), ‘Subject labels’ (e.g. *art* and *economics*) and ‘Regional labels’ (e.g. *Australian* and *mainly American*), whereas *Oxford Advanced Learner’s Dictionary of Current English* (OALD) makes no such division. In OALD, geographical restrictions are indicated by abbreviations such as *BrE* for British English and *WAfrE* for West African English and a separate list called ‘Labels’ covers other types of usage labels, everything from *technical* to *humorous*. *Collins*

*COBUILD Advanced Learner's English Dictionary* (COBUILD), then again, introduces two sets of labels: 'Geographical labels' and 'Style labels'. Labels denoting register (e.g. *business* and *medical*), style (e.g. *informal* and *formal*) and what Jackson (2002, 113) calls "effect" (e.g. *offensive* and *rude*) are all listed under 'Style labels'.

### 2.3.1 Formal words and the label *formal*

Formal words tend to be learned terms which have come to the English language from Latin and Greek (Jackson and Zé Amvela 2000, 138). The use of formal words has several reasons and consequences, of which Jackson and Zé Amvela (*ibid.*, pp. 138-139) mention the following three:

- 1) accuracy – "the precise technical names for ordinary language words"
- 2) appropriacy – "[t]he formal word has a distancing or euphemistic effect"
- 3) erudition – "a way of putting on airs or sounding posh or erudite"

It can thus be stated that some formal words may be useful when discussing, for example, scientific topics or a topic that might make people uncomfortable if more direct terms were used. Words and expressions such as *deceased*, *obese* and *visually impaired* function well in formal context where the distancing effect of the words is desired. However, using overly formal words in everyday English may have an unfriendly, ironic or humorous effect.

All the dictionaries used in this study (see Section 3.1.3) include the label *formal* and there are no major differences in their definitions of the label:

used mainly in official situations, or by political and business organizations, or when speaking or writing to people in authority (COBUILD, p. xii)

in current use but not used in ordinary conversation or in normal everyday writing (MEDAL, p. i)

usually only used in serious or official language and would not be appropriate in normal everyday conversation (OALD, p. i)

In addition to the definitions, the dictionaries give examples of formal words. In MEDAL the examples include *perpetrate* ('commit'), in OALD *admonish* ('chide'), and the only example given in COBUILD is *gratuity* ('donation', 'tip').

### 2.3.2 Informal words and the label *informal*

Some dictionaries use the term *colloquial* to refer to the words here called *informal* (Jackson and Zé Amvela 2000, 139). The term *informal* was chosen to be used in this thesis to create a more natural connection between the informal and the formal words. Furthermore, all three dictionaries consulted for this study have chosen to use the term *informal*.

According to Jackson and Zé Amvela (2000, 139), some informal words result from abbreviations (e.g. *ad* from *advertisement*) or coalescences (e.g. *gonna* from *goin to*) and others “are conventionally restricted to informal contexts” without a particular reason. All three dictionaries used agree on the definition of this stylistic label. According to them, informal words and expressions are used in informal and unofficial situations and would be considered inappropriate in formal contexts. In addition to the definitions, the dictionaries give examples of informal words. In MEDAL the examples include *guy* ('a man'), *bloke* ('a man') and *gutsy* ('brave and determined'), in OALD *bonkers* ('crazy') and *dodgy* ('dishonest', 'criminal'), and the only example given in COBUILD is *pep talk* ('encouraging speech or talk').

### 2.3.3 Slang words and the label *slang*

Of the three stylistic labels focused on in this thesis, *slang* is by far the most difficult to define. Compared to informal words, slang words can be said to occur in more informal situations (Jackson and Zé Amvela 2000, 139), they can be defined as “extremely informal”, “often fashionable and therefore rather temporary” (Andersson and Trudgill 1992, 16), or as informal words which have “not yet reached wide enough acceptance” to be labelled informal or colloquial (Jackson and Zé Amvela 2000, 139). The reason behind the use of slang is often social and it is used by the insiders of a certain

group to show who the outsiders are (Andersson and Trudgill 1992, 16, 79 and Battistella 2005, 86). All the definitions provided are fairly vague and it could be argued that of the labels discussed in this thesis, *slang* is most open to subjective interpretation. Jackson (2002, 112) points out that as a result of this, there is considerable variation between different dictionaries when it comes to labelling slang words.

The term *slang* is no longer used to mean the language of British criminals, as it once was (Andersson and Trudgill 1992, 77). One could argue that as the meaning of the term has changed, so have the functions of slang. The following list of four characteristics of slang is a summary by Battistella (2005, 85) of what was originally written by Dumas and Lighter (1978, 14-15) to help students in recognising true slang words which, in their view, should “meet at least two of the following criteria” (ibid., p.14):

- 1) Its presence will markedly lower, at least for the moment, the dignity of formal or serious speech or writing.
- 2) Its use implies the user’s special familiarity either with the referent or with that less satisfied or less responsible class of people who have such special familiarity and use the term.
- 3) It is a taboo term in ordinary discourse with persons of higher social status or greater responsibility.
- 4) It is used in place of the well-known conventional synonym, especially in order (a) to protect the user from the discomfort caused by the conventional item or (b) to protect the user from the discomfort or annoyance of further elaboration.

(Battistella 2005, 85)

The first characteristic has to do with the fact that slang is very low on the scale of formality. When slang is used, the context automatically changes into a less formal one. The second characteristic relates to the aforementioned fact that when using slang, one shows one’s membership in a social group. The third characteristic states that to avoid stigmatisation it is important to learn to recognise and use slang words in correct contexts and to refrain from using them with certain people and in certain situations. The fourth characteristic shows that despite differences in their level of formality, stylistically restricted words may have similar functions: slang words can be used as euphemisms in a way similar to formal words (see Section 2.3.1).



Not surprisingly, the definitions of the label *slang* differ in the three dictionaries examined. The list of stylistic labels used in COBUILD does not include this label. However, there are the labels *rude* and *offensive*, which might include some of the words labelled *slang* in other dictionaries. Other such words and expressions are probably labelled *informal* or *spoken* in COBUILD. However, this is only conjecture and the words found cannot be assumed to belong to the category of slang based on the labelling used in COBUILD. MEDAL does not use the label *slang* either, but it uses the label *very informal* instead. According to the definition, very informal words are “used only in very informal situations and mainly among people who know each other well. Some dictionaries use the label *slang*” (MEDAL, p. i). Based on the quotation, it is safe to assume that words labelled *very informal* in MEDAL can in fact be placed under the category of slang seen as there is the separate label *informal* in the dictionary. OALD is the only one of the three dictionaries to use the label *slang*, which is defined as follows: “*Slang* is very informal language, sometimes restricted to a particular group of people, for example people of the same age or those who have the same interests or do the same job” (OALD, p. i). As the clarifications of the labels *very informal* and *slang* in MEDAL and OALD show, the same notion is represented by different labels in these two dictionaries. In addition to the definitions, the dictionaries give examples of slang words. In MEDAL the examples include *go ape* (‘to become very excited or angry’), and *journo* (‘a journalist’), in OALD *dingbat* (‘a stupid person’) and *dosh* (‘money’).

## 2.4 Corpus linguistics, frequency and language teaching

Schoolbooks used in language teaching are often compiled using intuition-based methods trusting the innate knowledge the author has about the language and how it works (Biber and Reppen 2012a, xix-xx and Biber and Reppen 2012b, 239-240). However, O’Keeffe and Farr (2003, 390) point out that “numerous studies have shown that the language presented in textbooks is often based on faulty intuition about how people use language”. Sometimes the texts in schoolbooks are invented solely for the purpose of teaching grammatical constructions (Mishan 2004, 11-12 and Römer 2004, 153).

As an example of this, Carter (1998b, 45-49) takes dialogues in English language coursebooks and shows how they have been tidied up from the features characteristic of naturally occurring spoken language. Scrutinising the books used in teaching English as a foreign language in Germany, Römer (2004, 153) too, observed that the language was “simplified and non-authentic kind of English” and contained “sentences which probably have not occurred in any natural speech situation before (and which probably never will)”.

Corpus linguistics, an empirical and quantitative methodology to study how language is actually used in natural texts, predates the intuition-based methods which became popular in the 1960's and 1970's (Biber and Reppen 2012, xix-xxi and Biber et al. 1998, 11-21). Even as early as in the 18<sup>th</sup> century, collections of natural texts were used, for example, to identify meanings of words and thus served as basis for dictionary writing (ibid.). Thanks to the large electronic corpora available to us since the 1980's, using computer-based corpus linguistics as a research approach can provide us with reliable findings when studying lexicographical issues such as the commonness of a word, its collocations or its tendency to occur in a particular language variety (ibid.). Comparing the language in corpora with that in schoolbooks is another way of using electronic corpora and the findings from this type of schoolbook analysis can be used to improve schoolbook writing in a way that brings the natural language use and the language taught in schools closer to one another (Römer 2008, 115). Studies combining corpus and textbook analysis have been conducted to gain information on how, for example, specific grammar patterns or regional varieties are presented in schoolbooks (ibid.). However, the language of teaching resources has not yet been “systematically analysed” (Römer 2004, 151) and studies of different registers and varieties are still required (Römer 2008, 122-123). The present study answers to the need of comparative analyses of corpora and schoolbooks (ibid.) with the focus on stylistically marked words and levels of formality in the English language.

Schoolbook writing requires a wide understanding of the different pedagogical situations in which the materials might be used and what the needs of the users are (Bell and Gower 1998, 116). Introducing new vocabulary to learners is one of the important tasks of schoolbooks (Biber and

Reppen 2012b, 246) and McCarten (2007, 19-20) names two ways in which the books can help students to learn the words. According to her, the materials need to:

- present and practice in natural contexts the vocabulary that is frequent, current, and appropriate to learners' needs
- help students become better learners of vocabulary by teaching different techniques and strategies they can use to continue learning outside the classroom

(ibid.)

Biber and Reppen (2012b, 239-241) know that making decisions concerning the inclusion and exclusion of words requires careful judgment and they suggest that “[o]ne empirical basis that could be used for many of these decisions is frequency information”. They (ibid.), as well as Fox (1998, 27), come to the logical conclusion that the most common words would no doubt prove more useful than very infrequent ones, especially in the earlier years of language learning. Introducing rare words, then again, suites the purpose of widening the learner’s vocabulary but, as Biber and Reppen (2012b, 247) note, they should not occur in the books at the expense of the more common words. And, in my opinion, their restrictions should be made known to the language learner. In other words, if a word occurs very infrequently in natural language use, the schoolbook should provide the learner with the information on exactly how, where and when the word can be used and, furthermore, if there is a more common equivalent to it.

If a large corpus contains only a few instances of a word, the question arises as to whether it is reasonable to have the word occur in a schoolbook. Words like this “are usually more or less ignored by corpus linguists, as there is not enough evidence as to how they are typically used in the language” (Fox 1998, 26). Römer (2008, 115) too, argues for the inclusion of “the most common lexical items of a language and the patterns and meanings with which they typically occur” since these help learners to “develop both their receptive and productive skills” unlike the very infrequent ones. Indeed, several researchers have noticed that the language features which time and again cause problems among learners, are presented in schoolbooks in a way that has very little resemblance to authentic language use (see Römer 2008, 115). One way of conducting applied corpus research is to gather frequency

information from corpora and to compare it with the information gained by analysing schoolbooks (ibid.).

Learners may encounter difficulties even when learning very common words because of their various and often context-depended meanings (Fox 1998, 27). The different meanings are called word senses and listed in dictionaries one below another together with some information on their use. As for different words, frequency information can be used for word senses, too, and “more teaching and learning effort should go into the most frequent senses of words than into rare senses” (ibid.). Fox (ibid., p. 28) notes, however, that often a more concrete meaning of a word is taught even if the more abstract meaning occurs more frequently in natural language use. She suggests that this might happen because “it is easier to learn a concrete meaning than an abstract one” (ibid.). I agree with Fox (ibid.) and Römer (2008, 115) in that frequency information cannot be the only criterion for selecting words to include in schoolbooks, but it is such an important one that it should not be ignored.

### 3. Material and methods

The study was conducted in two stages: First, selected word lists (see Appendix I) in the chosen schoolbooks were searched for stylistically restricted words and the translations provided were compared to the dictionary meanings of the words. Second, the stylistically restricted words found were searched in corpora to discover their frequency of use. This section presents the material and methods in two subsections and follows the aforementioned order of the research. The first subsection describes how the schoolbooks were studied and provides information on the series of books and the dictionaries consulted. The second subsection focuses on the corpora and how they were utilised.

#### 3.1 Data collection and processing: Schoolbooks

The aim of this section is to explain the process of collecting the primary data for the present study. I will first introduce both the book series and the criteria used for choosing the series in question. I will proceed to describe the common themes in the books and how they affected the choice of the chapters and word lists studied. Before explaining the development of the analysis performed on the word lists, I will briefly introduce the dictionaries used to support the analysis.

##### 3.1.1 The three book series

For this study, I have chosen three series of schoolbooks used for teaching English in the last three grades of the Finnish basic education, namely the 7<sup>th</sup>, 8<sup>th</sup> and 9<sup>th</sup> grades. The series of books I will analyse are *Key English* and *Spotlight* from the publishing house SanomaPro (previously also known as WSOY and WSOYpro Oy) and *Smart Moves* from the publishing house Otava. These books were chosen for three main reasons: Firstly, all three book series are fairly recent and published after 2004; the year the current National Core Curriculum for Basic Education (NCCBE) was introduced. Secondly, because these books were published after 2004, they are compiled to follow the same NCCBE and contain similar themes. Thirdly, these series are widely used in Finland, hence, they create a representative sample of the books used for teaching English in Finnish secondary schools.

When referring to a certain book, I will add a number to the name of the series of books, for example, *Smart Moves 1*. The numbers 1, 2 and 3 refer to the grades 7, 8 and 9 respectively. This strategy was chosen over the original titles to keep the references to the books simple, to keep the numbering consistent, and to avoid confusion.

### 3.1.2 Themes and chapters

In this section, I will present the strategy on the basis of which the specific chapters and word lists were selected for the present study. The chapters are introduced below and their relation to the themes presented in the National Core Curriculum for Basic Education is also explained. A complete list of books, themes and chapters used for this study can be found in Table 1 below.

Table 1: Chapters by themes

Theme	Chapter Number	Chapter Name	Book series*
Immediate surroundings	1A	'Kids on the move'	<i>Key English 1</i>
	2	'The street where I live'	<i>Smart Moves 1</i>
	1A	'On the net'	<i>Spotlight 1</i>
School	3	'School is cool!'	<i>Key English 1</i>
	12	'School the Welsh way: Ysglo Gantaf'	<i>Smart Moves 1</i>
	3	'@school'	<i>Spotlight 1</i>
Travelling	10A	'Departures'	<i>Key English 2</i>
	17	'Land of shamrocks and leprechauns – excursion to Ireland'	<i>Smart Moves 2</i>
	4A	'The L.A. Street Lions'	<i>Spotlight 2</i>
Free Time and Hobbies	13B	'A Whole new ball game'	<i>Key English 2</i>
	1	'Aussie way of life'	<i>Smart Moves 2</i>
	2A	'Action weekend'	<i>Spotlight 2</i>
Media	26A	'The pros and cons of the media'	<i>Key English 3</i>
	23	'The internet – sex, lies and deceit'	<i>Smart Moves 3</i>
	3A	'Media mania'	<i>Spotlight 3</i>
Sustainable Development	27A	'Simple things you can do to help protect our environment'	<i>Key English 3</i>
	27B	'The state of the planet'	<i>Key English 3</i>
	4	'Global issues: energy'	<i>Smart Moves 3</i>
	5A	'Open your eyes'	<i>Spotlight 3</i>

\*1 = 7th grade

2 = 8th grade

3 = 9th grade

Altogether, the study comprises analysis of the word lists of eighteen chapters, that is, six chapters per each grade. To be able to compare the style of the words and their frequency of occurrence in the different series of schoolbooks and the different years of secondary school, I preferred examining chapters and texts which touch upon the same semantic field. Compared to random selection of chapters, this strategy increased the possibility of finding similar words or words of equal level of formality in the different series of books. Hence, I have chosen two themes for each grade and from each series of books I have selected one chapter per one theme. The final choice of chapters was made with the help of NCCBE and by comparing the contents of chapters in the different series of schoolbooks. According to NCCBE (Opetushallitus 2004, 141; my translation), the core contents that should be covered during secondary school are *free time and hobbies; travelling; public services; studying, work and business; sustainable development; health and well-being; and media*. In addition to these fairly broad themes, it is stated that the themes and topics familiar from primary school can also be covered.

Since each of the books used in the seventh grade begins with these revised themes, the first chapters to be looked at deal with the fairly simple theme of *Immediate Surroundings*, which is designed to provide an example on describing one's everyday life and living conditions. Since this is a theme already covered in primary school, it seemed logical to choose one of the first chapters where the repeated themes most likely appear. The decision was not based on the titles of the texts but on the thematic similarity to primary school material (i.e. descriptions of school, family and neighbourhood) which became apparent while reading the texts. The texts chosen are 'Kids on the move' (*Key English 1*), 'The street where I live' (*Smart Moves 1*) and 'On the net' (*Spotlight 1*), respectively numbered 1A, 2, and 1A in the books. The second theme chosen from the books used in the seventh grade was named *School* and the focus is on school systems and studying. The theme is familiar from primary school but recurs as *studying, work and business* in the core contents that should be covered during secondary school (Opetushallitus 2004, 139-141). The words were searched

in ‘School is cool!’ (*Key English 1*), ‘School the Welsh way: Ysglo Gantaf’ (*Smart Moves 1*), and ‘@school’ (*Spotlight 1*), which are marked in the books as units 3, 12 and 3 respectively.

From the seventh grade the study proceeds to the eight and towards the themes of *Travelling* and *Free Time and Hobbies* which present the vocabulary needed for life outside of home and school. For *Travelling*, the texts whose word lists were studied are ‘Departures’ (*Key English 2*), ‘Land of shamrocks and leprechauns – excursion to Ireland’ (*Smart Moves 2*) and ‘The L.A. Street Lions’ (*Spotlight 2*). The texts are numbered in the books as 10A, 17, and 4A respectively. The chapters chosen to represent the theme of *Free Time and Hobbies* are 13B: ‘A Whole new ball game’ (*Key English 2*), 1: ‘Aussie way of life’ (*Smart Moves 2*), and 2A: ‘Action weekend’ (*Spotlight 2*).

As the language skills develop towards the end of the secondary school, more specified and demanding themes are introduced to the students. The last themes to be covered in the scope of this study are *Media* and *Sustainable Development* – both present in the books for the ninth grade and both focusing on more societal, specific and current issues than the earlier themes. For the first theme, the most clearly media-related chapters in the schoolbooks are 26A: ‘The pros and cons of the media’ (*Key English 3*), 23: ‘The internet – sex, lies and deceit’ (*Smart Moves 3*) and 3A: ‘Media mania’ (*Spotlight 3*). For the second theme selected for the 9<sup>th</sup> grade, *Sustainable Development*, word lists of two relatively short texts were examined in the unit called ‘Environment’ in *Key English 3*: 27A ‘Simple things you can do to help protect our environment’ and 27B ‘The state of the planet’. The other word lists looked at belong to the texts 4: ‘Global issues: energy’ in *Smart Moves 3* and 5A: ‘Open your eyes’ in *Spotlight 3*.

### 3.1.3 Learner’s dictionaries

For the purpose of understanding the accurate meaning and level of formality of a word found in the word lists, I consulted *Oxford Advanced Learner’s Dictionary of Current English* (OALD), *Collins COBUILD Advanced Learner’s English Dictionary* (COBUILD) and *Macmillan English Dictionary for Advanced Learners* (MEDAL). The choice of dictionaries was made on the basis of how



comprehensive, respected and accessible they are. I chose to use learner's dictionaries since they are more likely to offer information on the style and usage of the words than general-purpose dictionaries. Although their lexical inventories are usually smaller than those of general-purpose dictionaries, learner's dictionaries can be expected to contain the majority of the words gathered from the schoolbooks, since both focus on (or should focus on) providing the learner with the most commonly used English words. To minimise the possibility of including words whose stylistically restricted nature is somehow controversial and to gain a wider understanding of the style and usage of the given word, I decided to use three dictionaries, instead of only one or two.

#### 3.1.4 Analysing word lists

The process of analysing the word lists of the chosen chapters was conducted in four stages. First, all the words that appeared to deviate from neutral style in any way were manually collected from the lists under examination. To increase the reliability of this stage, a fellow student was used as an inter-rater and rather too many than too few words and phrases were selected for further analysis. Phrasal verbs were, however, disregarded based on their frequent occurrence and the controversy in opinions on their use in formal and informal contexts (see Moon 2007, 915, Thim 2012, 197-246 and Wild 2011, 53-57). Second, all the words collected from the word lists were compared with the corresponding dictionary entries in the three learner's dictionaries presented in Section 3.1.3. In the third stage, the words that were labelled or described either as *formal*, *informal* or *slang* in two or more dictionaries were analysed further. Opinions on the level of formality of a word may vary, which makes labelling of words a complex task. To ensure that the stylistic label assigned by dictionaries is actually part of the connotation the word is commonly considered to carry, I decided to exclude from the analysis the words labelled in only one dictionary. Based on the labels, the words were categorised into formal, informal and slang words. In the fourth stage, the Finnish translation of the word was examined and evaluated against the information gained from the dictionaries. The target was to analyse the accuracy of the level of formality of the Finnish translation in comparison to the English

equivalent, in other words, whether the Finnish word is more formal or informal than the English word.

### 3.2 Data collection and processing: Corpora

This section provides information on how corpora were utilised in analysing the actual use of the words collected from the schoolbooks. Moreover, this section functions as an introduction to the corpora chosen as well as to the methods used in analysing the data gathered from the corpora. The main focus here is on the level of formality of the words, hence the *Corpus of Contemporary American English* (COCA) and its sub-corpora for spoken and written registers are introduced first. However, some of the words occur in texts that are clearly written to serve the purpose of introducing the learners to another variety of English, and hence the corpus of American English cannot be considered the most reliable source when analysing their use. As an informative addition to the formality of the words, the *Corpus of Web-Based Global English* (GloWbE) is used to show the distribution of the words in different varieties of English. Hence, the use of this corpus adds to the analysis of the natural context of the words and helps in examining whether the words actually represent the variety they have been connected to in the schoolbooks.

Both of the corpora were compiled by Mark Davies of Brigham Young University and accessed via the same online interface, which makes the comparison between different registers fairly effortless. Searches can be limited in several different ways, for example, by searching for words or lemmas or by tagging the search item. By searching for a lemma, the results will include all the realisations of that lemma i.e. all the inflected forms of the “dictionary form”. For example, the query for the exact word *make* will generate statistics and a list of examples for that word form only. However, placing square brackets around the search term changes it into a lemma search and *[make]* generates statistics and a list of examples which contain all the forms of the word *make*, hence the results will include, for example, the forms *makes*, *making* and *made*, too. Because the list of words gathered from the schoolbooks contains phrases too, in addition to the single words, the lemma search

run for each search item was disregarded as too complicated for the purposes of this study. Hence, the lemma search was used only in the cases where the word or phrase naturally (or only) occurred in some other than its basic form. On the other hand, tagging, which here means adding a part-of-speech tag (hereafter POS tags) to the search term, was used to some extent. With the help of the POS tags, words which have the same form but belong to different word classes could be separated from each other. This method was mainly used when searching the words in COCA where the actual frequency of occurrence was of more importance than when making comparisons between the geographical varieties in GloWbE. Therefore, the POS tags and how they were used is discussed in more detail in the section about COCA below.

### 3.2.1 COCA – *Corpus of Contemporary American English*

COCA was chosen to be used in the present study, firstly, because the size of it (more than 450 million words) is a significant factor when considering the reliability of the study. In the quotation below, Meyer (2002, 14-15) explains why the size of the corpus is vital for lexicographical studies.

While studies of grammatical constructions can be reliably conducted on corpora of varying length, to obtain valid information on vocabulary items, it is necessary to analyze corpora that are very large. To understand why this is the case, one need only investigate the frequency patterns of vocabulary in shorter corpora, such as the one-million-word LOB Corpus. In the LOB Corpus, the five most frequent lexical items are the function words *the*, *of*, *and*, *to*, and *a*. The five least frequent lexical items are not five single words but rather hundreds of different words that occur from ten to fifteen times each in the corpus. These words include numerous proper nouns as well as miscellaneous content words such as *alloy*, *beef*, and *bout*. These frequencies illustrate a simple fact about English vocabulary (or, for that matter, vocabulary patterns in any language): a relatively small number of words (function words) will occur with great frequency; a relatively large number of words (content words) will occur far less frequently.

The second reason lies already in the name of the corpus; it is a corpus of *contemporary* English. COCA covers texts from years 1990-2012 and it is a monitor corpus i.e. updated on regular basis (Davies, 2008b). It has been argued that only a corpus of this kind can be reliably used in studying lexical items due to the fact that meanings change and new words appear constantly (Meyer 2002, 15). This feature makes COCA suitable for the present study since the intention is to discover how the words are currently used. Because the level of formality of words has a tendency to change over

time, much older texts and conversations cannot be considered as an equally reliable source of information. The third reason for choosing the corpus in question is the convenient online interface and the possibility to search separately the sub-corpora of spoken, fiction, magazines, newspapers, and academic texts. These sub-corpora contain texts from various sources and are described as follows:

Spoken: (95 million words [95,385,672]) Transcripts of unscripted conversation from more than 150 different TV and radio programs (examples: *All Things Considered* (NPR), *Newshour* (PBS), *Good Morning America* (ABC), *Today Show* (NBC), *60 Minutes* (CBS), *Hannity and Colmes* (Fox), *Jerry Springer*, etc). [...]).

Fiction: (90 million words [90,344,134]) Short stories and plays from literary magazines, children's magazines, popular magazines, first chapters of first edition books 1990-present, and movie scripts.

Popular Magazines: (95 million words [95,564,706]) Nearly 100 different magazines, with a good mix (overall, and by year) between specific domains (news, health, home and gardening, women, financial, religion, sports, etc). A few examples are *Time*, *Men's Health*, *Good Housekeeping*, *Cosmopolitan*, *Fortune*, *Christian Century*, *Sports Illustrated*, etc.

Newspapers: (92 million words [91,680,966]) Ten newspapers from across the US, including: *USA Today*, *New York Times*, *Atlanta Journal Constitution*, *San Francisco Chronicle*, etc. In most cases, there is a good mix between different sections of the newspaper, such as local news, opinion, sports, financial, etc.

Academic Journals: (91 million words [91,044,778]) Nearly 100 different peer-reviewed journals. These were selected to cover the entire range of the Library of Congress classification system (e.g. a certain percentage from B (philosophy, psychology, religion), D (world history), K (education), T (technology), etc.), both overall and by number of words per year

(Davies, 2008c)

These five sub-corpora are helpful when finding out the text-type where the words most typically occur and higher rate of occurrence in one sub-corpus than another might indicate a specific level of formality. My hypothesis is that the words labelled *informal* have higher frequencies in the Spoken, Fiction and Popular Magazines corpora than in the corpora of Newspapers and Academic Journals, and that the reverse is true for words labelled *formal*. The sub-corpus for academic English is especially useful with formal words and the sub-corpus for spoken English with informal and slang words which might not occur in writing. However, very informal words and phrases are unlikely to

be found on the basis of “the fact that the people knew that they were on a national TV or radio program, and they therefore probably altered their speech accordingly” (Davies, 2008d).

Most of the words were searched in the corpus in the form they occurred in the word lists of the schoolbooks. POS tags were assigned where necessary, irrelevant articles were removed but the relevant articles were included in the search and sometimes even tagged to separate phrases from words and common nouns from proper nouns. The POS tag or tags that limited the search to a particular word sense were discovered heuristically. The list of POS tags and abbreviations used in the corpus and in this study can be found in Appendix II. Examples (1a) and (2a) below illustrate the undesirable results a basic word search can generate, while examples (1b) and (2b) show the desirable results which form the majority of the hits after tagging the search item.

(1a) <b>bit</b>	“I’ve owned every bit of it” (NEWS: AssocPress)
(1b) <b>a.[RR21] bit.[RR22]</b>	“It’s a bit like a really bad war” (NEWS: AssocPress)
(2a) <b>globetrotter</b>	“The Harlem Globetrotter has arrived” (NEWS: Chicago)
(2b) <b>globetrotter.[nn*]</b>	“Wagner was considered a globetrotter for his time” (ACAD: GeographRev)

In (1a) *bit* refers to a part of something and does not occur in the same sense as in the schoolbook. By adding the article to the word and tagging them as a single unit in (1b), the informal meaning ‘slightly, a little’ is separated from all the other results containing the word *bit*. The tag *RR* used in (1b) denotes ‘general adverb’, “[t]he first of the two digits indicates the number of words/tags in the sequence, and the second digit the position of each word within that sequence” (Davies, 2008e). In example (2a) *globetrotter* is part of the proper noun *Harlem Globetrotter* which is a name of a basketball team. It is clearly far apart from the meaning the word is used in in the schoolbook where it means a person who travels around the world. Tagging the word as a common noun in (2b), was an attempt to exclude all the proper nouns from the results. The number of proper nouns in the results decreased significantly but they could not be completely omitted by tagging the search word. Here, the frequency calculated by the corpus tool is much less reliable than it is with some other words. Excluding proper nouns from common nouns was one problem, but as mentioned in Section 2.4, a

word can have several different senses and these word senses can differ in how restricted they are in their use. One sense can be neutral and occur in various types of contexts and another sense can be restricted to a specific type of language use and these cannot be separated from each other by tagging. To find the instances of the word in the restricted sense, *globetrotter*, as well as other similar cases (*pest*: nuisance vs animal or insect, *Yankee*: an American vs part of the name of a baseball team, *Down Under*: Australia vs prepositions ‘down’ and ‘under’, *mad*: crazy vs angry, *box*: television vs container, and *jerk*: stupid or annoying person vs sudden movement or seasoning), a sample of the results was taken and the incorrectly tagged items were removed manually. The sample taken was either the complete list of results, if there were less than five hundred hits, or a random sample of five hundred hits. The approximate frequency for the latter case was then calculated using the following equation:

$$frequency = \frac{\text{instances counted manually}}{\text{size of the sample}} \times \text{instances counted by the corpus tool}$$

Example (3), using the word *globetrotter*, shows that for the small samples, the instances counted manually equal the frequency of the word in the corpus. Example (4), using the word *mad* (in the sense of “crazy”), shows how the approximate frequencies were calculated for words with five hundred hits or more.

(3) *globetrotter*

$$frequency = \frac{8}{20} \times 20$$

(4) *mad*

$$frequency = \frac{131}{500} \times 14,478$$

The corpus tool found twenty hits of the word *globetrotter* in COCA, eight of which in the meaning intended. Since there were 14,478 hits for the word *mad*, a random sample of five hundred was taken with the help of the corpus tool and the results were analysed to get an approximate frequency instead of manually analysing nearly 15,000 lines of text. In this case, 131 of the five hundred hits were considered relevant to the present study. The proportion gained by dividing the number of the relevant

hits by the sample size ( $131/500 = 0.262$ ) was used to discover how many of all the instances counted by the corpus tool could be considered relevant for the present study. This was conducted by multiplying the complete number of instances by the decimal number denoting the proportion ( $14,478 \times 0.262$ ) resulting in the approximate of 3,793.25 instances of *mad* in the sense of “crazy” in the whole corpus.

Of the special cases listed above, *box* in the less formal sense of ‘television’ (found in *Spotlight 3* in unit 3A: ‘Media mania’ and in *Key English 3* in chapter 26A: ‘The pros and cons of media’) produced difficulties. Tagging the word does not generate desirable results since the neutral sense and the restricted sense belong to the same part of speech. A synonym search was conducted on this word in a way shown in Figure 1 below.

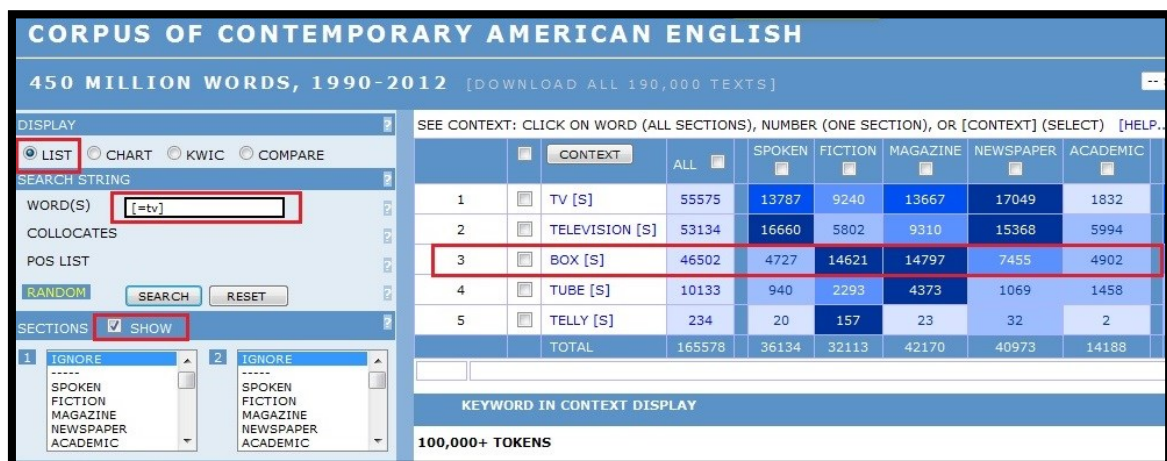


Figure 1: Synonym search for “tv” to find out the frequencies of the word “box” in the sub-corpora of COCA.

Since the word *box* was used in the schoolbook in an informal sense to mean ‘television’, a near synonym *tv* was used as a search term by placing it inside the square brackets after the equal sign. Settings were selected to generate a list of synonyms and to show the instances of the synonyms in every sub-corpora. It became clear from even a cursory examination of the results that not all the instances of *box* found using this method occurred in the sense of a television in the texts. Manual analysis of the random sample of five hundred instances of *box* in COCA revealed that none of the hits were in the restricted sense. Since the word was labelled *informal* it seemed apt to take a closer

look at the results in the ‘Spoken’ section of COCA. Another sample of five hundred instances of *box* but now in the sub-corpus for spoken language was searched manually to find all the instances where it occurred in the meaning ‘television’.

### 3.2.2 GloWbE – *Corpus of Web-Based Global English*

The *Corpus of Web-Based Global English* (GloWbE) is used in this study to gain an understanding of how certain words found in the schoolbooks are used in different varieties of English. GloWbE suits the purpose because it contains “1.9 billion words from 1.8 million web pages in 20 different English-speaking countries” (Davies, 2013). GloWbE consists of the language used on the internet and seems appropriate for this part of the study since the language used online tends to be more informal than in other contexts and therefore is more likely to contain regionally restricted words and expressions. With the help of the online interface the searches can be limited to chosen varieties. The chapters chosen from the schoolbooks contain references to the United States, Canada, Ireland, Australia and New Zealand, hence only these sub-corpora were searched. If necessary, the same search items and tags were used in GloWbE as in COCA (see Section 3.2.1).

### 3.2.3 Frequency calculations and organising the data

When a word is entered in the search field, the corpus tool counts the number of hits among the millions of words the corpus contains – the result is called the *raw frequency*. The raw frequency is rarely a very illustrative description of the relations between the occurrences of different words, especially when comparing the results from different corpora or sub-corpora which tend to have considerable differences in size. For example, the *United States* sub-corpus in GloWbE contains over 386 million words in contrast to the *Ghana* sub-corpus which contains over 38 million words. Both corpora are very large but the difference in size cannot be disregarded when analysing the results. As was pointed out in Section 3.2.1 above, content words can have low frequencies compared to function words, hence, percentages (i.e. occurrences per hundred words) would not be very illustrative when studying a very large corpus like COCA or GloWbE. To make the results comparable with each other,



the raw frequency is divided by the total number of words in the corpus and then multiplied by million as shown below.

$$\text{standardised frequency} = \frac{\text{raw frequency}}{\text{size of the corpus}} \times 1,000,000$$

This formula provides us with a *standardised frequency*, in other words, it shows how many times the word occurs per million words. The corpus tool calculates the standardised frequencies automatically but the formula was needed for the approximate frequencies (see Section 3.2.1). The size of the whole COCA was calculated from the sizes of the sub-corpora mentioned in Section 3.2.1 and the value used in the formula was the total number of words, namely 463,993,256.

The standardised frequencies of the words in the whole COCA and all the five sub-corpora were collected in Excel tables which were processed further to find out which words occurred most frequently in which sub-corpora. The words were arranged and new tables were created based on the frequency of occurrence, first, in the whole corpus to find out how common the words are in general, second, in the Spoken sub-corpus to see if some words tended to occur in informal context more often than others, and third, in the Academic Journals sub-corpus to examine which words are acceptable in formal language. The seven more problematic words, whose frequencies were not gained automatically (see Section 3.2.1), appear only in Table 7 (see Section 4.2.1) showing the frequencies of the words in the whole COCA since reliable frequency information from the sub-corpora was not available.

The words were categorised into three sets based on their frequencies. This strategy was chosen to be able to provide easily understandable information on the commonness of the word. The first category contains words which occur very frequently in natural language use. The second category contains little less common words and the third category holds the infrequent words. The limit for the very frequent words was gained by comparing the frequencies to the list of the five thousand most

frequent words in COCA (Davies, 2010). The frequencies in the list range from 5,079 up to 22,038,615. Based on this information, the minimum frequency for the first category was set to 5,000 instances and the maximum frequency of the third category was set to the tenth of this, i.e. in 500 instances. The second category contains the words whose frequency is between 500 and 5,000.

To understand why certain words occurred so infrequently in COCA the words in the third category were searched in GloWbE, too. This was done in order to find out whether these words were regionally restricted and occurred more frequently in some other variety than American English.

## 4. Results and analysis

In this section, I will present the results gained, first, by analysing the word lists of schoolbooks and, second, by searching the words in corpora in order to learn more about their restricted nature. I will show how the books differ by series and grades in the number of stylistically restricted words they contain and how the words found divide into the three categories of formal, informal and slang. The information provided by the books on the use of the words is also examined. The frequency information from the corpora is used in analysing the commonness of the words in authentic discourse. Comparison between the books is made based on the frequency levels of the words found in them. Graphs are used to demonstrate how large share of the words found in each book series is of which frequency level. In the last section, the results from the sub-corpora of COCA and GloWbE provide us with more precise information about the use of these words in different contexts.

### 4.1 Stylistically restricted words in schoolbooks

This section presents the findings of the analysis of the schoolbooks. The words found are categorized first by grade, then by style (*formal*, *informal* or *slang*) and finally by book series. After discussing the findings grade by grade, a summary of the findings is presented. For the complete list of the words found and their translations in the schoolbooks see Appendix III.

#### 4.1.1 Stylistically restricted words: 7<sup>th</sup> grade

The first theme examined in the schoolbooks used for teaching English as a foreign language in the seventh grade is *Immediate Surroundings*. The words that deviated from what could be considered as neutral style are all informal in nature. The following list includes the words which are labelled informal in at least two of the dictionaries and were found in the lists of words accompanying the three texts:

**Key English 1**

*a kid*  
*a city slicker*  
*dad*  
*a guy*  
*a hick*  
*buzz off*

**Smart Moves 1**

*pest*  
*goalie*

**Spotlight 1**

*you bet*  
*the States*  
*Howdy*

Words labelled informal in only one of the dictionaries were *mind you!*, *a bit* (*Key English 1*), *cute* (*Smart Moves 1*), *the net*, *lucky you!* and *cool* (*Spotlight 1*). Examples of other stylistically restricted words found in the schoolbooks are *a bumpkin* and *button his lips* (*Key English 1*). The first is labelled *disapproving* in OALD and COBUILD and described as insulting in MEDAL, the latter is described as spoken and rude in MEDAL. The words cannot be considered neutral but neither can they be placed under the categories of *formal*, *informal* or *slang* since none of the labels or descriptions given in the dictionaries has directly to do with formality.

The Finnish translations provided by the schoolbooks of the words listed above do not seem to keep in line with the original style of the words. For example, *city slicker* is translated as “kaupunkilainen, juppi” (meaning ‘a person living in a city, yuppie’), which seems inadequate since it should not be expected that the students of the target age group using the book are still familiar with the meaning of the word *juppi/yuppie* used in the translation. Another example is the translation “hei, moi” (meaning ‘hi’) of *Howdy*, where the humoristic tone of the English word is lost. In general, the Finnish translations seem to be a little higher on the scale of formality than the words they translate. Besides the translations, there are no other instructions on the use of these words.

The second theme chosen for scrutiny is *School*, which is a familiar theme already from the primary school but recurs as *studying*, *work and business* in the core contents that should be covered during secondary school (Opetushallitus 2004, 139-141). The words were searched in ‘School is cool!’ (*Key English 1*), ‘School the Welsh way: Ysglo Gantaf’ (*Smart Moves 1*), and ‘@school’ (*Spotlight 1*), which are marked in the books as units 3, 12 and 3 respectively. Again, all the stylistically restricted words found fall under the category *informal*. The lists below, give all the words which are labelled or described as *informal* in at least two dictionaries:

**Key English 1**

*crush*  
*pretty*  
*luv=love*  
*cussing*

**Smart Moves1**

*mate*

**Spotlight 1**

*how come*  
*laid back*

Three more words are labelled *informal* but only in one dictionary and are therefore listed separately here: *cool*, *cute* (*Key English 1*), and *tricky* (*Spotlight 1*). Similarly to the words listed previously in this section, the above words have become more formal when translated into Finnish. Of these words, *luv=love*, *mate*, and *how come* cause the most concern. Their translations in the schoolbooks are “terveisin” (often translated as ‘regards’, ‘greetings’ or ‘yours’), “kaveri” (meaning ‘friend’), and “miksi” (meaning ‘why’) respectively. None of the words is accompanied with a stylistic label, comment or anything that would indicate the stylistic restrictions of the word. Native speakers of English automatically recognise that there are only certain limited contexts where *luv* and *regards*, *mate* and *friend*, and *how come* and *why* are interchangeable. Using these words might reveal to the listener what is the level of formality of the context, which regional variant is represented, and/or what is the socio-economic status of the speaker. However, this information is not readily available to the foreign language learner and assistance is needed. *Mate* is often associated especially with Australian and British English and in the book it occurs in a text whose title, *School the Welsh way*, implies its connection to the British Isles. In the word list, however, there are no indicators on which of the words are regionally restricted and which can be more freely used.

The following table shows the number of occurrences of stylistically restricted words in the word lists examined in the books used in the 7<sup>th</sup> grade. The first figure shows the number of the words which are labelled in all three dictionaries. The bracketed figure shows the number of the words labelled in at least two of the dictionaries and the figure in the square brackets represents the number of words labelled in at least one of the dictionaries.

Table 2: Formal, informal and slang words: 7<sup>th</sup> grade

Theme	<i>Key English 1</i>			<i>Smart Moves 1</i>			<i>Spotlight 1</i>		
	F	I	S	F	I	S	F	I	S
Immediate Surroundings	-	5 (6) [8]	-	-	2 (2) [3]	-	-	2 (3) [6]	-
School	-	2 (4) [6]	-	-	1 (1) [1]	-	-	2 (3) [4]	-

F = Formal; I = Informal; S = Slang; 1 = occurrences, labelled in all the three dictionaries; (1) = occurrences, labelled in at least two dictionaries, [1] = occurrences, labelled in at least one dictionary

As can be seen from the table, stylistically restricted words are rare in the texts examined. Moreover, the difficulty of labelling the words is visible in the variation of the three successive figures.

#### 4.1.2 Stylistically restricted words: 8<sup>th</sup> grade

The choice of the first chapters for the eighth grade was based on the common theme of *Travelling* which is visible in both the topics of the texts and in the word lists containing words related to transportation and travelling in general.

The word lists studied included words and phrases which were considered to deviate from neutral style. The following words are labelled *informal* in at least two of the three learner's dictionaries:

##### ***Key English 2***

*a granny*

*a kid*

*gran*

*a stomach bug*

##### ***Smart Moves 2***

*crack a joke*

##### ***Spotlight 2***

*gonna = going to*

Information on the use of the words was attached only to the word *gonna* in *Spotlight 2*. The piece of advice provided was that its use should be restricted to spoken context only. Already this type of information can help students to pay more attention to the use of words and develop as language users. With the phrase *crack a joke* there has been an attempt to show that the word deviates from the neutral style. The phrase has been translated into Finnish as “kertoa, ‘murjaista’ vitsi” which seems fairly adequate but still seems to lack the colloquial touch the English word has. Unfortunately, the Finnish translation used in the book is not something a teenage-Finn would use in speech.

Phrases labelled informal in only one of the dictionaries are *sleep like a log* (*Smart Moves 2*) and *No sweat!* (*Spotlight 2*). The use of the first of these two phrases may not cause difficulties for students since it is used in a rather similar way in both languages. However, *No sweat!* is translated as “Helppoa!” (meaning ‘Easy!’ or ‘That’s easy!’) which only slightly implies informality with the exclamation mark.

*Knackered* is a word found in *Smart Moves 2* and categorised here as *slang* based on the labelling in MEDAL and OALD. However, the dictionaries are not unanimous in labelling the word and whereas COBUILD labels it *informal*, MEDAL and OALD place the word slightly further down on the scale and label it *very informal* and *slang* respectively. The word is translated as “poikki, ‘rättiväsynyt’”, which fairly well describe the meaning of the word but still lose some of the connotations attached to the English word. The translations do not make it explicit that *knackered* is a slang word nor that it is mainly used in British English. Again the problem with the translation seems to be in the difference of the level of formality: whereas the English word is categorised as *slang*, the Finnish translation is closer to *informal*.

The next chapters chosen to be examined belong to the second theme selected for the eighth grade, *Free Time & Hobbies*. The list below contains all the stylistically restricted words found which were labelled *informal* in at least two dictionaries.

***Key English 2***

*a Yankee*  
*a bit of*  
*a guy*  
*The States*

***Smart Moves 2***

*Down Under*  
*mad*  
*crazy*  
*I guess*  
*footy*  
*globetrotter*  
*Aussie*

***Spotlight 2***

*stuff*

The word *Yankee* is labelled *informal*, *often disapproving* in OALD, only *informal* in COBUILD, and described as an insulting word in MEDAL. The translation “jenkki” lacks the negative connotation attached to it in some varieties of English. Since the word is not marked as being informal and/or insulting, the students may learn the word only partially and use it inappropriately in

the future. For the next phrase, *a bit of*, the translation “melkoinen, aikamoinen” (meaning ‘considerable’ or ‘quite a’) is appropriate, but this sense of the phrase seems to be both informal and fairly rare, too. OALD labels the word *informal* and describes it as especially British, COBUILD does not even include this sense of the phrase, and MEDAL labels it *informal*. Based on the inconsistency of the findings, it could be concluded that the use of the phrase is indeed restricted and some information about its natural context and collocates might be useful.

The word *guy* is translated fairly successfully as “mies, kaveri, tyyppi” (meaning ‘man’, ‘friend’, ‘guy’ or ‘dude’). The three words tell the learner that the word is used to refer to a male friend in an informal situation. The word “tyyppi” would have been enough to convey the informality but since it is gender-neutral, the other translations are useful, too. On the other hand, the word is translatable with only one word because there are many informal Finnish words denoting the same thing.

*The States*, *Down Under* and *Aussie* have been provided with rather inadequate translations. “*Yhdysvallat*” means ‘United States’ or ‘United States of America’, “*Australia (ja Uusi-Seelanti)*” means ‘Australia (and New-Zealand)’ and “*australialainen*” means ‘Australian’. None of the translations implies the informality of the words even though all three dictionaries unanimously label these words *informal*. The rest of the words listed above, i.e. *mad*, *crazy*, *I guess*, *footy*, *globetrotter*, and *stuff* are less likely to be used in a wrong context. However, a greater accuracy in the translations or a labelling of some kind might be useful here, too. For example, OALD differentiates *mad* and *crazy* by informing the reader that the first is mainly used in British English and the latter in North American English. The translation of *I guess* is marked as being limited to the text or context in which it appears in the book by adding “tässä:” to the actual translation “varmaankin” which means ‘in this text: probably’. This is a good way of raising the language learner’s interest, but it may also leave the learner uncertain of the type of contexts where the word can actually be used. The next word, *footy*, is interesting because in all three dictionaries (sometimes spelled *footie*) it is described as *informal* and as being a British English word for football but in the schoolbook it is translated as



“australialainen jalkapallo”, which means ‘Australian football’ – a sport closer to American football and rugby than what is conventionally meant by football in Europe. Here, the same note “tässä”, ‘in this text’, could be used to limit the translation of the word to the context (free time in Australia) if the connection to British English seems irrelevant for the text in question. However, the fact that none of the dictionaries mention Australian football in connection with the word *footy* raises the question whether Australians in fact use the word and if this actually is part of the meaning of the word. The last two words, *globetrotter* and *stuff*, are labelled *informal* in all three dictionaries. They are accurately translated as “maailmanmatkaaja” and “jutut, asiat” and the use of the words should not create difficulties for the learner.

The words labelled *informal* in only one dictionary are:

**Key English 2**

*a whole new ball game*  
*dozens of*

**Smart Moves 2**

*cute*  
*G'day*  
*Take care!*

**Spotlight 2**

*youngster*  
*you must be joking*

These words and phrases have been translated relatively well in the books and, for example, the idiomatic nature of *a whole new ball game* and *you must be joking* is made explicit by translating them as a whole instead of translating them word by word. The translation provided for *a whole new ball game* is “myös: aivan uusi asia”, which means ‘also: a completely new thing’ and *you must be joking* is translated as “et voi olla tosissasi”, ‘you cannot be serious’. *Dozens of*, *cute* and *youngster* are translated as “kymmeniä, suuri määrä”, “söpö” and “nuori” respectively and are fairly unambiguous even though the latter lacks the old-fashioned tone, which, according to MEDAL, accompanies the English word. *G'day* is described as an Australian way saying ‘hi’ and the exclamation mark in the translation “pärjäilkää!” for *Take care!* implies informality and the spoken context.

The following table shows the number of occurrences of stylistically restricted words in the word lists examined in the books used in the 8<sup>th</sup> grade. The figures marked with an asterisk refer to

the word *knackered*, which is labelled *informal* in one of the dictionaries and *slang* in the two others as mentioned previously in this section.

Table 3: Formal, informal and slang words: 8<sup>th</sup> grade

Theme	<i>Key English 2</i>			<i>Smart Moves 2</i>			<i>Spotlight 2</i>		
	F	I	S	F	I	S	F	I	S
Travelling	-	4 (4) [4]	-	-	0 (1) [2/3*]	0 (1*) [1*]	-	1 (1) [2]	-
Free Time & Hobbies	-	2 (4) [6]	-	-	6 (7) [10]	-	-	1 (1) [3]	-

F = Formal; I = Informal; S = Slang; 1 = occurrences, labelled in all the three dictionaries; (1) = occurrences, labelled in at least two dictionaries, [1] = occurrences, labelled in at least one dictionary

Table 3 shows how few words deviating from neutral style the texts contain. The fact that there are no formal words in the first three chapters examined is surprising. The theme of travelling would have enabled the use of a more versatile vocabulary. For example, formal words could be introduced through such situations as making reservations and purchasing tickets. Not surprisingly, the chapters about free time and hobbies include more informal words than the first three chapters. Interesting is also the difference in the number of occurrences between the three books.

#### 4.1.3 Stylistically restricted words: 9<sup>th</sup> grade

The first theme examined in the books used in teaching the last grade of the secondary school is *Media*. The words listed below were found in the texts representing this theme and are labelled *informal* in at least two dictionaries:

##### ***Key English 3***

*the box*  
*an airhead*  
*a pal*  
*a kid sister*  
*a jerk*  
*No way!*  
*an ad*

##### ***Smart Moves 3***

*wacko*  
*ad=advert*  
*=advertisement*  
*porn*

##### ***Spotlight 3***

*a telly*  
*a box*  
*on the go*

Most of the translations provided do not imply the informality of the English words, for example *wacko* is translated as “hullu” which is closer to ‘crazy’ in its meaning and level of formality.

Nevertheless, one could assume that students would not use words such as *box* (meaning ‘TV’), *airhead*, *jerk*, *wacko*, *porn*, or *telly* in too formal situations. Likewise, the phrases *No way!* and *on the go* cannot be interpreted as anything but idiomatic and hence restricted in their use. The words which students may understand and use incorrectly are *pal*, *kid sister* and *ad* respectively translated as “kaveri” (meaning ‘friend’), “pikkusisko” (meaning ‘little sister’) and “mainos” (meaning ‘advertisement’). *Smart Moves 3* provides three forms for the word *advertisement* and thus shows students how the shorter words have been formed. This might be enough to imply the informality of the words *ad* and *advert*. In *Key English 3*, however, there is no extra information given of the word *ad* and the user of the book is left alone to learn about the style of the word. For the word *kid sister*, there may not even be any Finnish equivalent implying the level of informality but for the word *pal* there are plenty. Therefore, it is surprising to notice how formal a translation has been chosen.

Words labelled *informal* in one dictionary only are:

<b><i>Key English 3</i></b>	<b><i>Smart Moves 3</i></b>
<i>rubbish</i>	<i>spamming</i>
<i>corny</i>	<i>spam</i>

The translations “roska(a)” and “typerä, tylsä” (provided for *rubbish* and *corny* respectively) clearly imply that the meaning of the words is such that they should not be used in a formal context, but especially the translation for *corny* could be more informative since the context of the English word is more restricted than the context of the Finnish words. *Spamming* and *spam* are difficult words to define stylistically because they belong to the field of computing and are naturally more contextually restricted than the other words listed above.

The first formal words found in this study appeared when examining the chapters under this theme. There were two instances in *Smart Moves 3* which were labelled *formal* in all three dictionaries, namely *thus* and *acquire*. Quite an accurate translation has been provided for the word *thus*, which is translated into Finnish as “täten” in the word list. The Finnish word is almost exclusively used in writing or in formal spoken situations and, therefore, even the translation is enough to help students understand the stylistic restrictions of the English word. Contrary to this, the

translation “saada haltuunsa” given for the word *acquire* does not imply much of the level of formality or the preferred context of use. Since the word is clearly considered *formal* in English, it might be justifiable to expect some notes to appear on its use. In addition to *thus* and *acquire*, another slightly formal word was found in *Spotlight 3*: the word *dislike* is labelled *rather formal* in OALD and translated as “ei pitää” (word for word: ‘not like’) in the book. Since the translation does not convey the stylistic restrictions, it might be clarifying to offer, for example, parallel but stylistically differing example sentences: *dislike* in the more formal sentence and an informal equivalent in the less formal one. If this is impossible due to limitations of space, marking the words somehow might be considered.

The second theme selected for the 9<sup>th</sup> grade is *Sustainable Development*. As might have been expected, the stylistically restricted words found in these chapters are all labelled *formal*. What is surprising is the small number of these words: *appliance* and *vehicle* (*Key English 3*), which are labelled *formal* in one dictionary, and *vegetation* (*Spotlight 3*), which is labelled in two of the dictionaries. In addition to these words, there were several which the dictionaries describe as belonging to a certain field of science but since these words are not labelled *formal*, they cannot be taken into account in this study. The translations “kulkuväline” for *vehicle* and “kasvillisuus” for *vegetation* are unambiguous and seem to be in accordance with the style of the English word, only the translation “laite” for the word *appliance* might be considered slightly less formal.

The table below shows the number of occurrences of stylistically restricted words in the word lists examined in the books used in the 9<sup>th</sup> grade.

Table 4: Formal, informal and slang words: 9<sup>th</sup> grade

Theme	<i>Key English 3</i>			<i>Smart Moves 3</i>			<i>Spotlight 3</i>		
	F	I	S	F	I	S	F	I	S
Media	-	5 (7) [9]	-	2 (2) [2]	3 (3) [5]	-	0 (0) [1]	3 (3) [3]	-
Sustainable Development	0 (0) [2]	-	-	-	-	-	0 (1) [1]	-	-

F = Formal; I = Informal; S = Slang; 1 = occurrences, labelled in all the three dictionaries; (1) = occurrences, labelled in at least two dictionaries, [1] = occurrences, labelled in at least one dictionary

As can be seen from the figures, the number of stylistically restricted words does not increase as the studies proceed but there is slightly more variation in the styles found in the texts (for comparison see Table 2 and Table 3).

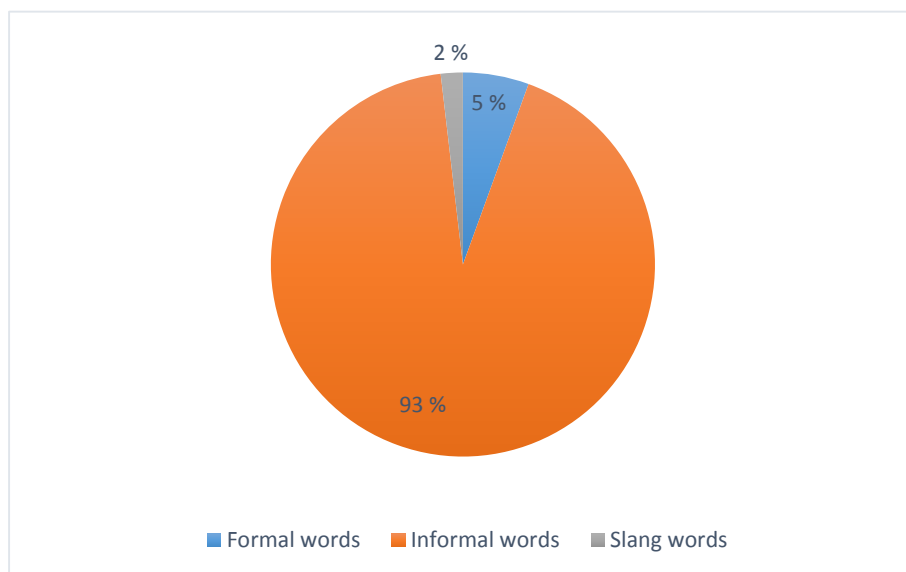
#### 4.1.4 Summary of findings

Altogether the study concerned eighteen chapters, whose word lists contained 2,035 words in total, of which seventy-nine (3.9%) I considered stylistically restricted and analysed further. Of these seventy-nine words, fifty-four were labelled *formal*, *informal* or *slang* in at least two of the three learner's dictionaries consulted. In this section, the findings are presented graphically to show how the words found divide into the three stylistic categories and how frequent such words are in different book series and grades. Table 5 shows the number of formal, informal and slang words, as well as the total of stylistically restricted words, found in each chapter. Only the words labelled *formal*, *informal* or *slang* in at least two dictionaries are included in the figures to minimise the possibility of including words whose stylistically restricted nature is somehow controversial. Table 5 also shows the total number of words in each word list and the percentage of stylistically restricted words compared to the total.

Table 5: Summary of the number of words found

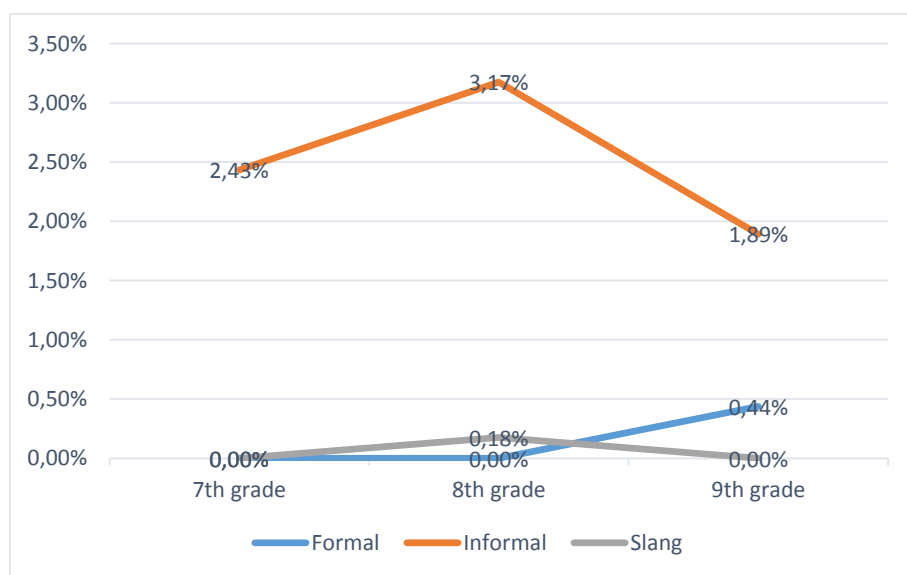
Grade	Theme	Book	Formal words	Informal words	Slang words	Stylistically restricted words (total)	Words (total)	Percentage
7	I	<i>Key English 1</i>	0	6	0	6	164	3.7 %
7	II	<i>Key English 1</i>	0	4	0	4	218	1.8 %
8	I	<i>Key English 2</i>	0	4	0	4	88	4.5 %
8	II	<i>Key English 2</i>	0	4	0	4	121	3.3 %
9	I	<i>Key English 3</i>	0	7	0	7	113	6.2 %
9	II	<i>Key English 3</i>	0	0	0	0	128	0.0 %
7	I	<i>Smart Moves 1</i>	0	2	0	2	69	2.9 %
7	II	<i>Smart Moves 1</i>	0	1	0	1	98	1.0 %
8	I	<i>Smart Moves 2</i>	0	1	1	2	75	2.7 %
8	II	<i>Smart Moves 2</i>	0	7	0	7	92	7.6 %
9	I	<i>Smart Moves 3</i>	2	3	0	5	112	4.5 %
9	II	<i>Smart Moves 3</i>	0	0	0	0	74	0.0 %
7	I	<i>Spotlight 1</i>	0	3	0	3	89	3.4 %
7	II	<i>Spotlight 1</i>	0	3	0	3	143	2.1 %
8	I	<i>Spotlight 2</i>	0	1	0	1	90	1.1 %
8	II	<i>Spotlight 2</i>	0	1	0	1	101	1.0 %
9	I	<i>Spotlight 3</i>	0	3	0	3	131	2.3 %
9	II	<i>Spotlight 3</i>	1	0	0	1	129	0.8 %
<b>Total</b>			3	50	1	54	2035	2.7 %

The graphs presented in this section are compiled based on the figures in Table 5 above. As mentioned in Section 4.2 there are more words labelled *informal* than *formal* in dictionaries and as expected the pattern is the same in schoolbooks. The following graph shows how the fifty-four stylistically restricted words found divide into the three categories *formal*, *informal*, and *slang*.



Graph 1: Division of stylistically restricted words into formal, informal and slang

Graph 1 shows that the great majority of stylistically restricted words in schoolbooks used for teaching English in Finnish secondary schools are indeed informal in nature. The share of formal words is rather small and slang words are in clear minority, accounting for only two per cent of all the stylistically restricted words found. Even though the share of informal words remains prevalent in each of the word lists, it does not stay constant. Graph 2 below shows how the share of formal, informal and slang words changes when the students move from one grade to the next.

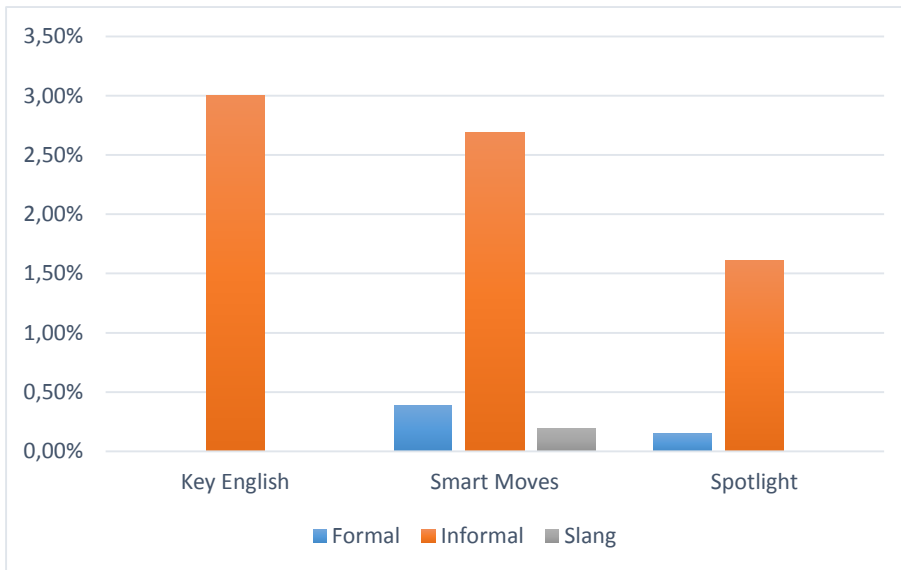


Graph 2: Stylistically restricted words in grades seven, eight and nine

It can be interpreted from Graph 2 that the share of informal words increases in the word lists when students move on to the eighth grade. In the ninth grade, the share of informal words decreases and is lower than on the seventh grade. The share of the slang words found follow the same pattern but on a smaller scale: the only occurrences of words labelled *slang* are on the eighth grade. As for the formal words, the pattern is different: the share of formal words in the word lists increases from zero close to 0.5 per cent on the ninth grade. The word lists on the seventh and eighth grade do not include any formal words. The pattern described could be explained by the tendency of formal words to be more difficult to learn, as they are often Latin-based and longer than informal or slang words (see Section 4.2.1) and hence they are introduced later. Another contributory factor might be the themes which affect the style of the language in the chapters: the themes are clearly more scientific in nature in the ninth grade than in the two preceding grades and hence more formal vocabulary might be expected. For the style of a text to be considered constant, the increase in formal vocabulary naturally decreases the number of informal and slang words. On the other hand, the increase in informal vocabulary lowers the formality of a text and thus allows more slang words to occur. Surprisingly, this pattern cannot be found in the word list of Theme I in *Smart Moves 3*. There are two formal and three informal words in the same word list.

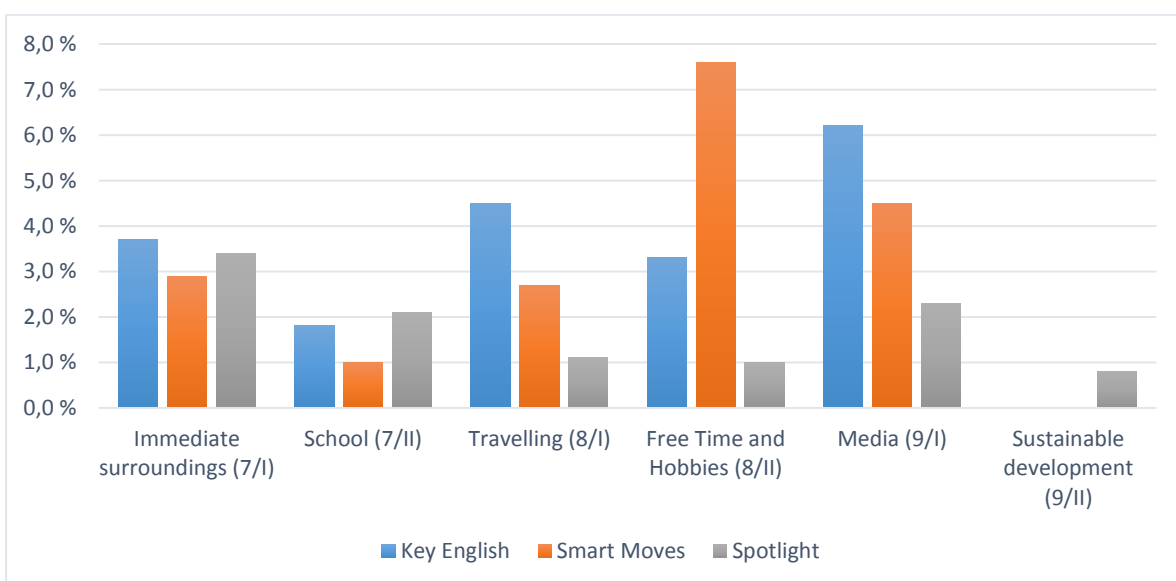
To be able to compare the occurrence of stylistically restricted words in the three series of books, the chapters and their word lists examined were chosen based on semantic similarity (see Section 2). However, semantic similarity did not result in stylistic similarity. Graph 3 below shows how the series of books differ in their use of stylistically restricted words.





Graph 3: Share of formal, informal, and slang words in the three book series

Graph 3 presents the percentage of stylistically restricted words compared to the total word count of the six chapters that were looked at in each of the series of books. Of the three series of books the most versatile vocabulary is found in *Smart Moves*, which is the only series where words belonging to all three stylistic categories were found. The occurrence of stylistically restricted words appears to be the lowest in *Spotlight* and the least variation is found in *Key English*.



Graph 4: Comparison of book series by the number of stylistically restricted words in each theme 7, 8, 9 = the grades; I = the first theme; II = the second theme

Graph 4 illustrates how the share of stylistically restricted words varies between the chosen themes in the three book series. The reason behind the decision to choose the word lists for analysis from chapters and texts which touch upon the same semantic field, was made to increase the possibility of finding similar words or words of equal level of formality in the different series of books. Graph 4 shows that this strategy resulted in fairly equal percentages in three of the six themes. *Immediate surroundings*, *School* and *Sustainable development* contain relatively similar amount of stylistically restricted vocabulary in all of the three books, but greater differences between the book series can be seen to occur with word lists dealing with vocabulary for *Travelling*, *Freetime and hobbies* and *Media*.

## 4.2 Formal, informal and slang words in corpora

This section presents the findings gained by studying the words discussed above in Section 4.1 in COCA and GloWbE. The focus is on how frequent the words are in general to analyse if and how they should be taught in school. The numbers presented on their occurrence in the different sub-corpora can, on the one hand, be used as a guideline for their use in correct context and, on the other hand, function as a tool for evaluating whether the words should be marked in schoolbooks.

### 4.2.1 Frequencies of the words in COCA

Of the stylistically marked words found in the schoolbooks, the most frequent word in COCA was found to be *guy* with 77,105 hits and the standardised frequency of 166.05 per million. The least frequent words were *box* (in the sense of ‘television’) with no hits and *luv* (used at the end of a letter or e-mail in stead of ‘love’) with three hits and the standardised frequency of 0.01. Because of the vast differences in the frequencies, the words were divided into three categories. The first category contains two out of the three (66.67%) formal words and twelve out of the 47 (25.53%) informal words. All these words occurred over 5,000 times in COCA and their standardised frequencies range from 12.09 to 166.05 instances per million words. The third formal word belongs to the second

category which contains all the words whose frequency is between 500 and 5,000. Similarly to the first category, twelve out of the 47 (25.53%) informal words belong to the second category. Here, the raw frequencies are between 710 and 4,014, and the standardised frequencies between 1.53 and 8.64 instances per million words. Considering that the words examined are from schoolbooks used for teaching English as a second language in Finnish secondary schools, it is somewhat surprising that the third category contains more words than the first two categories with the more common words. Altogether, there are 24 words in the third category out of which 23 are informal and one is a slang word. Almost half (48.94%) of the informal words found in the schoolbooks occur less frequently than once in a million words and have less than 500 instances in the corpus of over 450 million words. The frequency information of the words is presented in Table 6, Table 7 and Table 8 below. In the tables, an asterisk is placed before the words whose frequencies could not be gained directly from the corpus tool but are generalisations (rounded to two decimals) based on manually analysed samples from the corpus (see Section 3.2.1). Names of the books are shortened in the table into KE (*Key English*), SM (*Smart Moves*) and SL (*Spotlight*).

Table 6: Frequencies of the formal words in COCA

FORMAL				
	WORD	BOOK/THEME	PER MILLION	FREQUENCY
1. CATEGORY (FREQ > 5000)				
1	thus	SM3/I	135.17	62,764.00
2	acquire	SM3/I	13.30	6,176.00
2. CATEGORY (5000 > FREQ > 500)				
3	vegetation	SL3/II	8.64	4,014.00

Table 7: Frequencies of the informal words in COCA

INFORMAL				
	WORD	BOOK/THEME	PER MILLION	FREQUENCY
1. CATEGORY (FREQ > 5000)				
1	guy	KE1/I, KE2/II	166.05	77,105.00
2	pretty	KE1/II	144.69	67,187.00
3	*stuff	SL2/II	101.55	47,152.00
4	dad	KE1/I	88.02	40,873.00
5	I guess	SM2/II	83.49	38,769.00

6	kid	KE1/I, KE2/I	77.65	36,058.00
7	a bit	KE1/I	56.76	26,354.00
8	gonna	SL2/I	55.67	25,850.00
9	crazy	SM2/II	51.15	23,752.00
10	an ad	KE3/I, SM3/1	25.99	12,068.00
11	a bit of	KE2/II	21.64	10,049.00
12	mate	SM1/II	12.09	5,615.00
<b>2. CATEGORY (5000 &gt; FREQ &gt; 500)</b>				
13	*mad	SM2/II	8.18	3,793.24
14	a pal	KE3/I	7.44	3,453.00
15	how come	SL1/II	6.04	2,806.00
16	the States	SL1/II, KE2/II	5.63	2,613.00
17	crush	KE1/II	4.51	2,092.00
18	You bet	SL1/I	4.46	2,069.00
19	a granny	KE2/I	3.94	1,830.00
20	porn	SM3/I	3.74	1,735.00
21	*jerk	KE3/I	2.83	1,312.70
22	gran	KE2/I	2.12	984.00
23	Mind you!	KE1/I	1.60	742.00
24	goalie	SM1/I	1.53	710.00
<b>3. CATEGORY (FREQ &lt; 500)</b>				
25	howdy	SL1/I	0.93	430.00
26	on the go	SL3/I	0.92	425.00
27	Aussie	SM2/II	0.75	349.00
28	No way!	KE3/I	0.64	296.00
29	*Yankee	KE2/II	0.63	292.57
30	laid back	SL1/II	0.62	288.00
31	a telly	SL3/I	0.50	233.00
32	wacko	SM3/I	0.49	228.00
33	*Down Under	SM2/II	0.44	205.94
34	hick	KE1/I	0.37	174.00
35	*pest	SM1/I	0.26	118.65
36	a kid sister	KE3/I	0.25	115.00
37	[crack] a joke	SM2/1	0.14	64.00
38	buzz off	KE1/I	0.14	63.00
39	an airhead	KE3/1	0.14	63.00
40	advert	SM3/I	0.11	49.00
41	city slicker	KE1/I	0.08	37.00
42	cussing	KE1/II	0.08	37.00
43	a stomach bug	KE2/I	0.03	12.00
44	*globetrotter	SM2/II	0.02	8.00
45	footy	SM2/2	0.01	5.00
46	luv=love	KE1/II, SL1/II	0.01	3.00
47	*box	KE3/I, SL3/I	0.00	0.00

Table 8: Frequency of the slang word in COCA

SLANG				
	WORD	BOOK/THEME	PER MILLION	FREQUENCY
3. CATEGORY (FREQ < 500)				
1	knackered	SM2/I	0.04	18.00

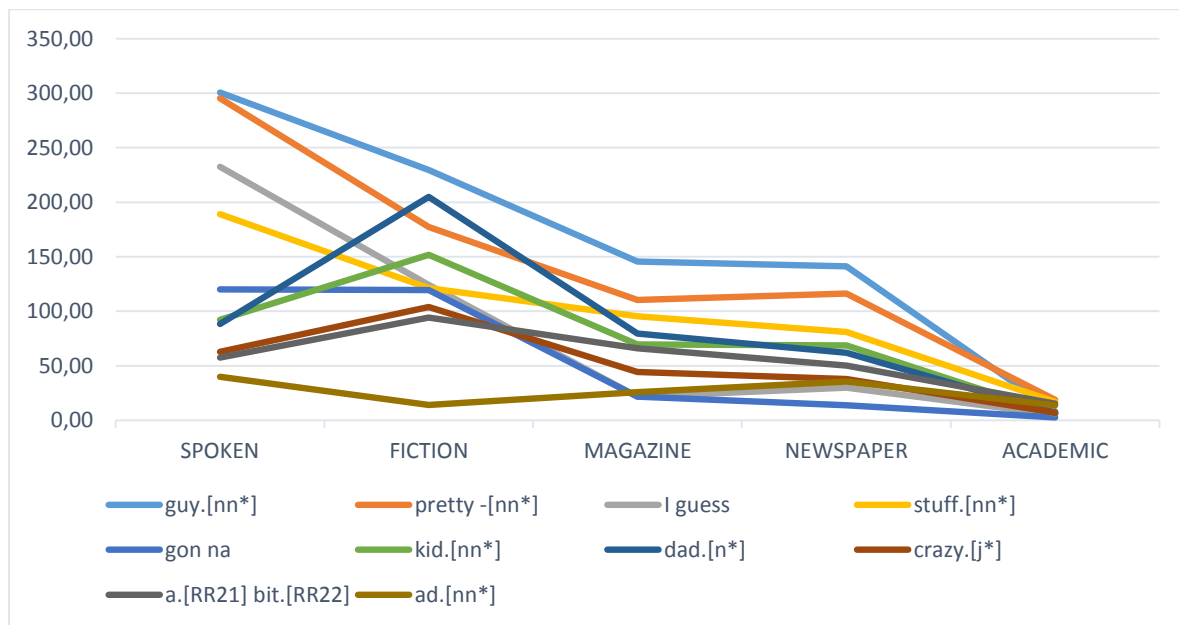
#### 4.2.2 Distribution of the frequent words in sub-corpora

Of the informal words, the ten most frequent in the Spoken and in the Academic sub-corpora are listed below together with their standardised frequencies. The complete table with the raw and standardised frequencies of all the words in each sub-corpora can be found in Appendix IV.

SPOKEN		ACADEMIC	
1. guy	300.67	1. pretty	18.92
2. pretty	295.44	2. stuff	17.03
3. I guess	232.43	3. a bit	15.43
4. stuff	189.17	4. ad	13.84
5. gonna	119.96	5. guy	8.26
6. kid	92.29	6. crazy	7.08
7. dad	88.41	7. the States	6.92
8. crazy	62.87	8. dad	6.84
9. a bit	57.57	9. mate	6.74
10. ad	39.91	10. kid	6.17

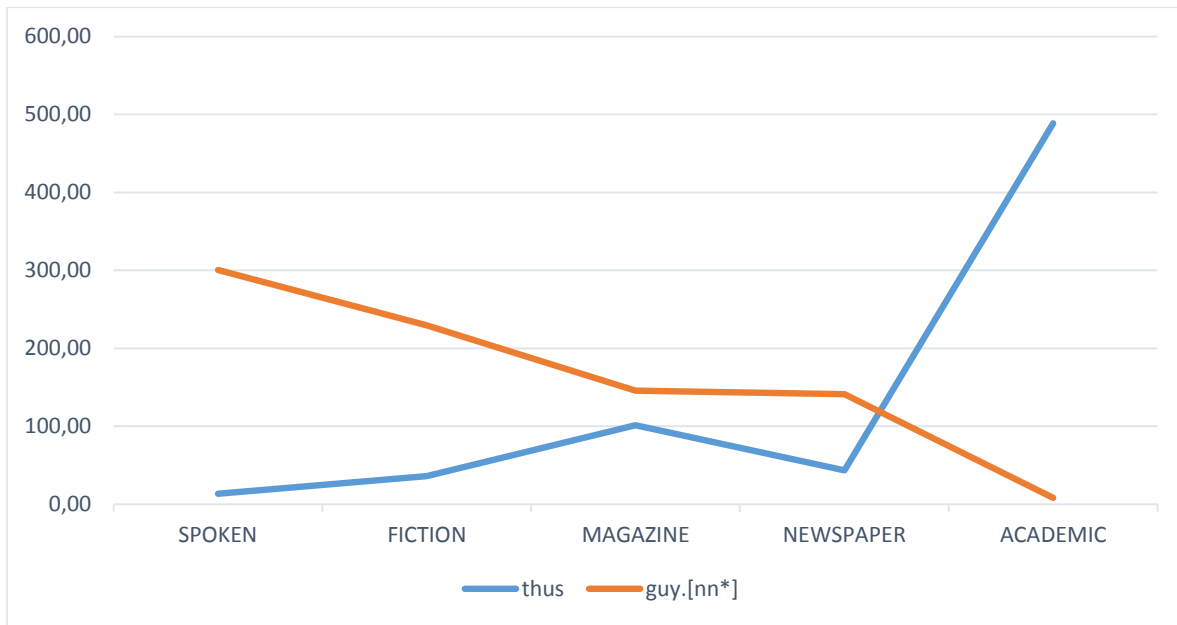
Comparison of the standardised frequencies of the two lists, confirms the informal nature of the words; the frequencies are much higher in the Spoken sub-corpus which represents informal language than in the Academic sub-corpus which represents formal language. Interestingly, the lists differ less than could have been expected. The only words present in the ACADEMIC –list but missing from the SPOKEN –list are *the States* and *mate*. Their relatively high occurrence rates might result from mistagging and some of the instances of *the States* are probably instances of the plural form of the word ‘state’ instead of the United States of America. *Mate*, on the other hand, might be incorrectly tagged as a noun instead of a verb. The high position of *guy* in the ACADEMIC –list might as well result from a failure to tag the proper noun *Guy*. The words missing from the ACADEMIC –list are *I guess* and *gonna* which are commonly known to be informal and mainly restricted to spoken

language. Graph 5 below illustrates how the words in the SPOKEN –list are represented in the five sub-corpora. The information from the graph can also be extended to depict how the words align on the scale of formality; the spoken and the academic sub-corpora representing the extremes of the scale where fiction, magazine and newspaper language appear in the order mentioned.



Graph 5: Ten most frequent words in the Spoken sub-corpus

Graph 5 shows that the frequencies of the words decrease gradually as the context becomes more formal. The most dramatic drop the frequency levels of *I guess* and *gonna* between FICTION and MAGAZINE. Their relatively high occurrence rate in FICTION can be explained by the dialogue and imitation of spoken language often present in fiction writing. Graph 6 below illustrates the different nature of the very informal and very formal words and how they are spread across COCA. The most frequent words in SPOKEN and ACADEMIC were chosen as representatives.

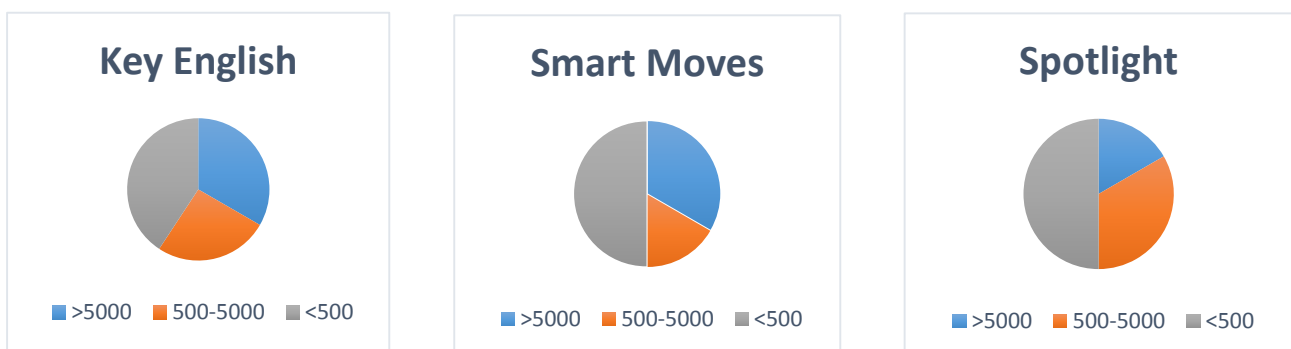


Graph 6: *Thus* and *guy* in the five sub-corpora

In Graph 6, *thus* in blue represents formal words and *guy* in orange represents informal words. The graph clearly shows how differently these words are used and how their acceptability depends on the context. The frequency of informal words decreases towards the more formal contexts and the frequency of formal words increases.

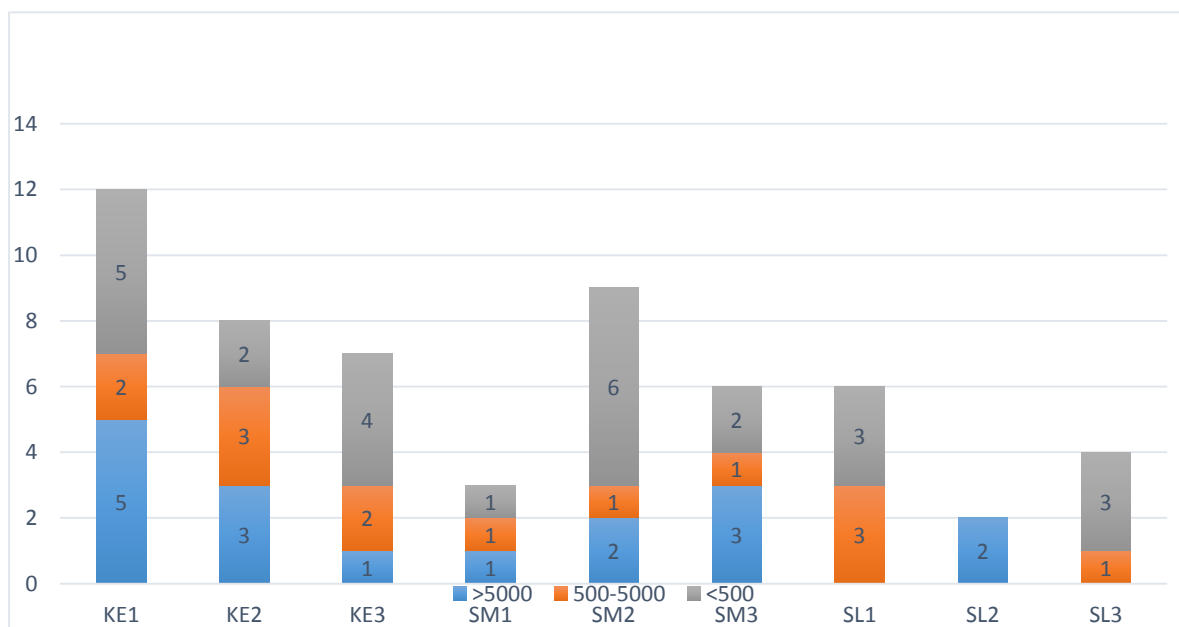
#### 4.2.3 Distribution of frequent and infrequent words in schoolbooks

Different schoolbooks have a very different number of stylistically restricted words in their word lists and the proportions of formal, informal and slang words was found to vary, too (see Section 4.1.4). Comparison of the books regarding the frequency levels of the words, shows that the book series differ also in this respect. These differences are presented graphically in Graph 7 below.



Graph 7: Frequency levels of stylistically restricted words in the book series

Of the 27 words analysed in the series *Key English*, eleven belong to the third frequency category introduced in Section 4.2.1 and marked in grey in Graph 7 above. Hence, the very infrequent words cover 40.74 per cent of the stylistically restricted words found. The share is even higher in *Smart Moves* and *Spotlight*, in which the word counts for these words are nine and six respectively, resulting in 50 per cent in both the series. The smallest share of words belonging to the first category with over 5,000 occurrences in COCA is found in *Spotlight* books where only two out of the twelve words belong to this group. This 16.67 % share is shown in blue in Graph 7. Of the words found in *Key English* and *Smart Moves*, a third (33.33%) is very frequently occurring words. The percentage of the words in the second category is 25.93 in *Key English*, 16.67 in *Smart Moves* and 33.33 in *Spotlight*, meaning seven, three and four words respectively. This category contains the words whose frequency is between 500 and 5,000, and is represented in the graph by orange colour. The distribution of the words into the three frequency categories in each book can be seen in Graph 8 below.



Graph 8: Frequency levels of stylistically restricted words in each book  
 KE = *Key English*; SM = *Smart Moves*; SL = *Spotlight*

Similarly to Graph 7, in Graph 8, too, the grey colour denotes the less frequent words, the orange colour is for the middle category and the blue colour for the most frequent words. The common words



cover less than a half of the stylistically restricted words in all the books except for *Spotlight 2* and *Smart Moves 3*.

#### 4.2.4 Infrequent words in GloWbE

The words which had very low frequencies in COCA were searched in GloWbE to find out whether they were regionally restricted and therefore occurred less often in the corpus of American English.

The raw frequencies and the standardised frequencies gathered in GloWbE are presented in Table 9.

Table 9: Low-frequency words in GloWbE

WORD	ALL	US	CA	GB	IE	AU	NZ	ALL	US	CA	GB	IE	AU	NZ
howdy	713	267	47	111	24	53	19	0.38	<b>0.69</b>	0.35	0.29	0.24	0.36	0.23
on the go	5,518	840	523	1,288	324	441	361	2.93	2.17	3.88	3.32	3.21	2.98	<b>4.44</b>
<u>Aussie</u>	13,009	628	249	1,534	336	7,156	1,148	6.90	1.62	1.85	3.96	3.33	<b>48.28</b>	14.10
No way!	1,129	298	66	174	44	110	48	0.60	<b>0.77</b>	0.49	0.45	0.44	0.74	0.59
laid back	3,471	516	243	719	220	294	193	1.84	1.33	1.80	1.85	2.18	1.98	<b>2.37</b>
<u>telly</u>	4,517	191	53	2,638	562	355	218	2.40	0.49	0.39	<b>6.81</b>	5.56	2.40	2.68
wacko	531	324	30	60	7	42	13	0.28	<b>0.84</b>	0.22	0.15	0.07	0.28	0.16
<u>* Down Under</u>	4,045	453	157	817	254	959	367	2.15	1.17	1.16	2.11	2.51	<b>6.47</b>	4.51
hick	266	125	15	34	11	31	14	0.14	<b>0.32</b>	0.11	0.09	0.11	0.21	0.17
kid sister	128	44	4	13	7	3	3	0.07	<b>0.11</b>	0.03	0.03	0.07	0.02	0.04
buzz off	136	31	15	30	11	8	6	0.07	0.08	<b>0.11</b>	0.08	<b>0.11</b>	0.05	0.07
crack a joke	245	46	19	55	11	8	6	0.13	0.12	<b>0.14</b>	<b>0.14</b>	0.11	0.05	0.07
airhead	293	90	12	69	8	30	26	0.16	0.23	0.09	0.18	0.08	0.20	<b>0.32</b>
<u>advert</u>	7,617	334	94	4,000	516	249	260	4.04	0.86	0.70	<b>10.32</b>	5.11	1.68	3.19
city slicker	112	28	10	18	10	11	3	0.06	0.07	0.07	0.05	<b>0.10</b>	0.07	0.04
cussing	111	54	4	10	1	4	4	0.06	<b>0.14</b>	0.03	0.03	0.01	0.03	0.05
stomach bug	243	36	8	63	13	27	19	0.13	0.09	0.06	0.16	0.13	0.18	<b>0.23</b>
<u>* globetrotter</u>	115	19	8	26	3	7	6	0.06	0.05	0.06	<b>0.07</b>	0.03	0.05	<b>0.07</b>
<u>footy</u>	1,774	22	17	297	21	1,282	80	0.94	0.06	0.13	0.77	0.21	<b>8.65</b>	0.98
luv	8	0	1	1	0	1	1	0.00	0.00	0.01	0.00	0.00	0.01	0.01
<b>** box, collocate: [watch].[v*]</b>	343	53	15	134	29	37	10	0.18	0.14	0.11	<b>0.35</b>	0.29	0.25	0.12
knackered	1,152	31	14	739	92	77	64	0.61	0.08	0.10	<b>1.91</b>	0.91	0.52	0.79

All = Entire GloWbE; US = American English; CA = Canadian English; GB = British English; IE = Irish English; AU = Australian English; NZ = New Zealand English

In Table 9 above, the words which are clearly specific to a certain variety are separated from the rest of the words by underlining them. The highest standardized frequency for each word is bolded. Using

the highest standardized frequency as criterion, the words divide into the geographical varieties as follows:

US	CA	GB	IE	AU	NZ
<i>howdy</i>	<i>buzz off</i>	<u><i>telly</i></u>	<i>buzz off</i>	<u><i>Aussie</i></u>	<u><i>on the go</i></u>
<i>No way!</i>	<i>crack a joke</i>	<i>crack a joke</i>	<i>city slicker</i>	<u><i>Down Under</i></u>	<u><i>laid back</i></u>
<i>wacko</i>		<u><i>advert</i></u>		<u><i>footy</i></u>	<i>airhead</i>
<i>hick</i>		<i>globetrotter</i>			<i>stomach bug</i>
<i>kid sister</i>		<i>box</i>			<i>globetrotter</i>
<i>cussing</i>		<u><i>knackered</i></u>			

The words whose highest standardized frequency is over one per million are underlined in the lists above. Minority of the words being underlined, it can be stated that most of the words are uncommon also globally. *Aussie*, *telly*, *Down Under*, *advert* and *footy* have considerably higher frequency levels on certain regional varieties than in others as is shown in the bar graphs generated by the corpus tool in Figure 2-6 below.

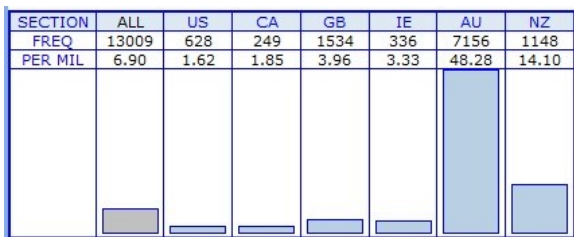


Figure 2: *Aussie* in GloWbE

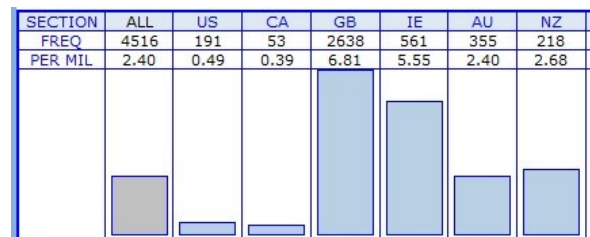


Figure 3: *Telly* in GloWbE

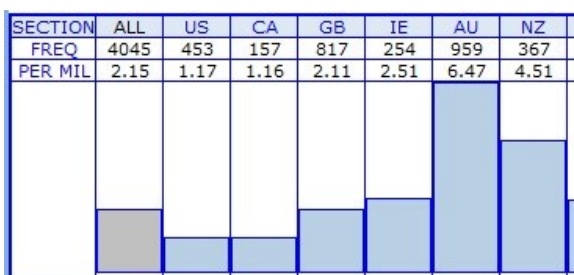


Figure 4: *Down Under* in GloWbE

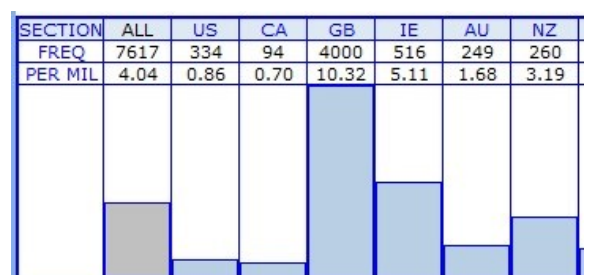


Figure 5: *Advert* in GloWbE

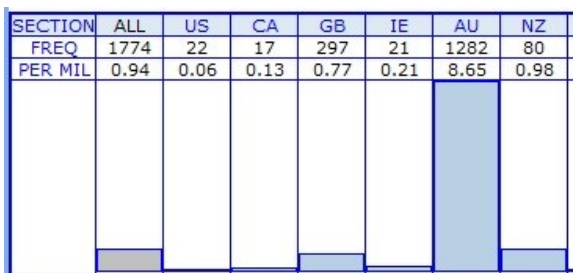


Figure 6: *Footy* in GloWbE

Even if all the five words above seem to be regionally restricted, the meanings of the words show that *Aussie* and *Down Under* are rather words related to the culture of the area than words which would have equivalents in other regional varieties. Similar results can be gained by searching *kangaroo* and *koala*, for example. Therefore, *Aussie* and *Down Under* are very different from the words *advert* and *telly* which are informally used in British and Irish English for the words *advertisement* and *television*. The reason for their low frequencies in COCA lies in their regionally restricted nature; they are not American English words and thus cannot be found in corpus for American English. *Footy* has the highest frequencies in the texts from Australia, New Zealand and Great Britain, but neither of the last mentioned gain standardized frequency above one. In the sub-corpus for Australian English, on the other hand, *footy* occurs 8.65 times in million words. Collocation lists reveal, however, that in Australian English the word often occurs in connection with *AFL*, short for Australian Football League, which refers to Australian rules football, not the European football, which is called *soccer* in American English and sometimes referred to as *footy* in British English. The examples (5) and (6) below demonstrate the different meanings of the word *footy* in the two varieties.

(3) “I feel like every time I start to get confident and start to play good **footy** in the AFL I was just getting injured again. [...]”

(AU G: news.com.au)

(4) “games, with EA’s rival FIFA franchise now dominating the living rooms of **footy** fans. But, the PES Productions team is not giving up without a fight”

(GB G: spong.com)

In (5) *AFL* refers to Australian rules football as mentioned above, and in (6) references to EA’s console games *FIFA*, Fédération Internationale de Football Association, and *PES*, Pro Evolution Soccer, relate the word *footy* to the European football. Interestingly, none of the dictionaries listed the meaning of *footy* exemplified in (5), but they did make a reference to the meaning in (6) (see Section 4.1.2). Based on the results from GloWbE, ‘Australian rules football’ is part of the meaning and should be mentioned in dictionaries since this meaning is more common than the British informal

use of the word. In this respect, *Smart Moves 2* has a more accurate description of the word than the dictionaries.

#### 4.2.5 Summary of findings

Searching different corpora for the stylistically marked words found in schoolbooks (see Section 4.1) revealed valuable information about their commonness in natural language use and about their stylistic and regional restrictions. The frequency levels between different words were found to vary drastically. Whereas the formal words were either very frequent or semi-frequent, the highest number of informal words divided into the category of very infrequently occurring words. The informal words with high frequencies are probably very familiar to most of the learners of English; words such as *guy*, *pretty*, *stuff*, *dad* and *gonna* are used in colloquial English worldwide. Moving further down on the list of frequencies (see Table 6), towards the very infrequently occurring *hick*, words with more restricted contexts are found: *You bet*, *jerk*, *howdy* and *wacko*, to name a few. Occasionally, the low frequencies in COCA could be explained by the findings from GloWbe. For example, *knackered*, the only slang word found, occurred very infrequent in COCA but similarly to many of the informal words, its frequency levels increased significantly when searched in other varieties of English. In COCA it occurred 0.04 times per million words whereas in GloWbe's sub-corpus for British English its standardised frequency was 1.91.

The results obtained by examining how word frequencies vary between the spoken, fiction, magazine, newspaper and academic sub-corpora of COCA, agree with the style labels assigned to the words in dictionaries and show to which contexts the use of the words is restricted. In a graphic description of standardised frequencies of a word in the five different sub-corpora, a steep decline from spoken to academic indicates informality, and vice versa, higher frequencies in the corpora for written English imply formality. These tendencies are shown in Graph 5 and Graph 6 in Section 4.2.2.

A significant inconsistency in the use of frequent and infrequent words with stylistic restrictions was discovered to exist both between and within chapters and book series (see Section 4.2.3). A single

chapter could contain very infrequent, semi-frequent as well as very frequent words. Comparing the book series revealed that *Key English* had nearly equal shares of words in all the frequency categories while the words in *Smart Moves* tended to be either very frequent or very infrequent. The share of the middle category i.e. semi-frequent words was higher in *Spotlight* than in *Smart Moves*, but very frequent words were scarce in *Spotlight*.

## 5. Discussion

In this section, I will return to the research questions asked in Introduction (see Section 1) and discuss the possible answers to them on the basis of previous literature on the subject (see Section 2 and Section 3) and the findings of the present study (see Section 4). I will also evaluate the limitations of the study and discuss the possible implications of the findings.

### 5.1 Answering the research questions

The study described in this thesis was conducted to discover how different styles are represented in language teaching materials. In the previous sections, I have presented the materials, methods, data and results of the study. The purpose of the following three subsections is to answer the research questions which initiated the process. The three main research questions are addressed separately in their own subsections below, the question functioning as a heading to the section.

#### 5.1.1 Do schoolbooks contain stylistically restricted words?

Based on the analysis of the chapter-specific word lists in the schoolbooks, the hypothesis (see Section 1) on the occurrence of stylistically restricted i.e. formal, informal and slang words is proven to be correct. As expected, there are stylistically restricted words in the books. Surprisingly, the proportion of these words is relatively small; only 2.7 per cent of the words can safely be said to belong to one of the three stylistic categories. In other words, 97.3 per cent of the words in the word lists can be considered to be neutral in style. Bearing in mind the various context in which learners use English, or will use later in life, the imbalance between the percentages seems vast. For the purpose of teaching how to use or how to recognise different styles, the share of stylistically restricted words in schoolbooks looks inadequate. In this respect, the books do not seem to support the national objectives set in NCCBE (see Section 2.1).

Considering the claim made earlier in this thesis about the language used by learners to be too formal (see Section 2.2), the lack of formal words in all three books is somewhat unexpected. The

dominance of informal words within the stylistically restricted words found (see Graph 1 in Section 4.1.4) does not provide support to the claim either. On the other hand, the majority of the words found being neutral, the learners might not encounter enough informal vocabulary to adjust their language for the informal everyday situations where even the neutral can appear too formal. Whether the learners using the book have the information needed on the formality of the words available to them, is discussed in the subsection below.

#### 5.1.2 Does their restricted nature become clear to the user of the book?

One possible explanation for learners blending different styles or using words in incorrect contexts (see Section 1 and Section 2.2) lies in the nature of word lists and the translations provided in them. Analysing the translations shows how insufficient the word lists are; the Finnish translation is often a tad more formal than the English word or phrase and there is nothing else in the lists to imply the stylistic restrictions of the words (see Section 4.1). This might not cause problems for advanced learners who are aware of the different levels of formality and the limitations of word lists, hence, are able to evaluate critically the information provided and be more cautious when using new words. However, most of the students in secondary school are only learning critical thinking and they trust the printed word, in other words, they cannot be expected to question the translations in the word lists. Moreover, based on the scarcity of stylistically restricted words in the schoolbooks used in language teaching (see Section 4.1), there is no reason to expect the students to be very much aware of the different levels of formality. All things considered, the word lists in schoolbooks might require more attention than is conventionally given to them – both when creating and using them.

Several researchers stress the importance of context for vocabulary learning (see Section 2.2) and there are plenty of examples on the possible pedagogical applications of corpora in the literature published in recent years (see Section 2.4). Nevertheless, little of this is visible in the schoolbooks examined in this study; the context in which the word occurs can sometimes be limited to a single heading and the patterns familiar from the texts are repeated in the exercises. The following

subsection suggests reasons for the restricted context and discusses the commonness and usefulness of the words found by referring back to the results gained from corpora (see Section 4.2).

### 5.1.3 How widely are the words used in authentic language?

The infrequency of some words cannot always be explained by the rarity of the concept or topic itself. Sometimes the frequency of a word is affected by restrictions on its use. It seems unlikely that a word occurs only a few times in a large corpus and each time in a completely different context. Furthermore, even if a word of this kind was found in a corpus, it should not be included in a schoolbook due to its unpredictable nature which would prevent anyone from learning the word given the fact that context is an integral part of the word's meaning.

The frequency of occurrence in corpora indicates how widely used a word is in authentic language and therefore it has commonly been used as a yardstick for the usefulness of the word (see Section 2.4). The corpus findings of the present study show that the majority of the stylistically restricted words in schoolbooks occur very infrequently in American English (see Section 4.2.1) and regional restrictions explain the absence of only a few of the words (see Section 4.2.4). The usefulness of these words for a secondary school student can be questioned and their inclusion in the books should be reassessed. Introducing very rare and stylistically restricted words in schoolbooks seems to support the theory that the language in them is to some extent based on incorrect assumptions on language use (see Section 2.4). In the defence of the schoolbook writers, something of the difficulty of choice of inclusion and translating words says the discovery that dictionaries are not always unanimous on the use of words (see Section 3.1 and Section 4.1) and the descriptions in dictionaries might not always be supported by the corpora findings as was seen with the word *footy* (see Section 4.2.1 and Section 4.2.4) which was described in the dictionaries as being a British English word for football but searching the word in GloWbe showed it to be used far more often in Australian English for Australian rules football.



## 5.2 Limitations and implications

In this section, I discuss the matters which might have had an effect on the results and reliability of the study. I also give suggestions on how the findings can benefit the field of language teaching and what should still be studied to be able to prepare learners for stylistic variation in the target language.

### 5.2.1 Limitations of the study

The restricted scope of the study suggests some caveats when it comes to the results. The schoolbooks were not examined from cover to cover but a sample of chapters was chosen to exemplify the situation. Different choice of chapters (or a more extensive study covering all the word lists) might have produced slightly different results. However, the number of stylistically restricted words in the word lists is low in each of the eighteen chapters, hence it can be stated with confidence that stylistically restricted words are not common in schoolbooks and the style of the books is mainly neutral.

Subjectivity in assigning style labels can be considered as another factor affecting the reliability of the results to some extent. The methods used (see Section 3.1.3 and Section 3.1.4) minimised the possibility of including words whose stylistic nature is controversial but resulted in excluding some words which would rarely be used in neutral context, for example, insulting words which have gained some other type of label in the dictionaries instead of *informal* or *slang* or words labelled *formal*, *informal* or *slang* in only one of the dictionaries. Hence the number of neutral words might be slightly different if a different criteria was used.

The frequency count of some of the words in corpora is an estimate rather than an accurate number due to the polysemous nature of the words and the consequential difficulty to generate adequately restricted searches. The estimated frequencies are calculated based on a random sample of five hundred instances of the word in corpora, which undeniably affects the reproducibility of the results. To increase the reliability of this section of the study, the same method was used with all the words causing similar difficulties and a fairly comprehensive sample size was used. A detailed description of the methods can be found in Section 3.2.1

### 5.2.2 Implications of the results and tasks for the future

The first step in reducing the gap between the authentic language use and the language taught in classrooms is to indicate where the two differ the most. Several writers have noted that learners encounter problems when trying to adapt their style of speaking or writing to the context (see Section 2.2). According to the findings of the present study, difficulties in finding the correct level of formality can be explained by the lack of attention to stylistically restricted words in schoolbooks and by the fact that words of different levels of formality are sometimes introduced within the same text and word list. Neither can be disregarded the finding that schoolbooks contain words which are very infrequent in authentic communication and of whose stylistic restrictions there is not much information available. The second step towards teaching authentic language and the various styles needed is to create materials which support these objectives. Whether creating, compiling or developing traditional schoolbooks, electronic books or online learning environments, teaching about stylistic variation should be included in the syllabus to better prepare the learners of English to the world outside the classroom. The possibilities range from marking the stylistically restricted words in the traditional word lists (and, for example, gathering the marked words to a table together with words of similar meaning but differing level of formality) to virtual learning environments with links and paths to corpora or extra audio or video examples. Cooperation between different agents (such as teachers, materials writers and researchers) as well as further research in the field is required. I agree with Römer (2008, 122-123) in that we need to discover the most problematic areas of learner language to be able to compile teaching materials which answer to learners' needs.

The importance of teaching stylistic variation for foreign language proficiency has been established several times in previous literature (see Section 2.2) but as Römer (2008, 122) claims, “we still have a long way to go in providing more adequate descriptions of different types of language (different text types, registers and varieties), based on larger collections of data”. For example, searching different corpora for the most commonly occurring formal and informal words and their collocations would be an interesting task for the future and important for materials development.

Moreover, research on the global language of social media (which could be referred to as modern slang) would be beneficial not only for the learners growing up with it, but also for the linguists and teachers trying to keep up with the changing language and the new ways of communication.

## 6. Conclusion

In this thesis, I have shown by analysing word lists of eighteen chapters in three book series how different styles of English are represented in foreign language teaching materials. One of the findings of the study is that schoolbooks contain stylistically marked words some of which can clearly be classified as being either formal, informal or slang. *Pal*, *thus*, *knackered* and *city slicker* are all words exemplifying the varying levels of formality found in the vocabulary introduced to secondary school students in Finland. The number of similar words with stylistic restrictions varies between and within the series of books. As an example of this, the highest combined percentage of formal, informal and slang words in a word list of a single chapter was 7.6 and the lowest was zero. Most of the words in the lists analysed were not formal, informal or slang based on the labels given to them in learner's dictionaries. The majority of the words in schoolbooks can be said to be neutral, even when taking into consideration the words with attitude labels (e.g. *offensive* or *rude*) and the words which were labelled as stylistically restricted in only one dictionary and hence excluded in this study. This result is not surprising, since the vocabulary taught at this level of study still contains relatively many very basic words. Informal words clearly outnumber both formal and slang words which are extremely rare in the books. This seems positive considering the learners' need to be prepared for colloquial conversations which make a large portion of everyday communication in English. Negative is the lack of information given on the use of the words. How can the learners know where they should use these words if the translations are too formal and the context provided is insufficient? Learning the words would require more examples of their use in authentic language. *Wacko*, *porn*, *thus* and *acquire* exemplify the problem of including both formal and informal words in one and the same word list. The sentences where they occur in the text might guide on their use, but their restrictions should be shown when listing the words by separating them from each other or by marking them differently. Another matter of concern is the frequency of the words: Most of the words found in the books are very infrequent in natural language use and some of their frequencies vary regionally.

The findings indicate that schoolbooks pay too little attention on an important matter. The concept of style should not be overlooked in foreign language teaching, since it can have a great effect on how the speaker or writer is perceived by others. The stylistic differences and levels of formality should at least be understood by the learner, even if adapting their own language use to the situation still requires more practice. The low frequency levels of the words used in schoolbooks, found in corpora, suggest that the language in the teaching materials is not very authentic and that intuition-based methods might still be used in schoolbook writing.

From the point of view of learners and teachers, the findings of the study are significant. If no improvements are made to schoolbooks, more work and effort is required in foreign language classes to learn the correct use of the words. Moreover, to meet the objectives set in the NCCBE, teachers in Finland should not trust blindly on the books but take responsibility in teaching the students to become aware of the differences between the different variants of English. The results of the study show that there are many possibilities for materials development in the future concerning the teaching of stylistic matters – and increasing authenticity in the books is not the least considering its proven effect on learners' motivation.

## References

### Dictionaries

Rundell, Michael, ed. 2007. *Macmillan English Dictionary for Advanced Learners*. 2<sup>nd</sup> edition. 1<sup>st</sup> edition 2002. Oxford: Macmillan Publishers Limited.

Sinclair, John, ed. 2006. *Collins COBUILD Advanced Learner's English Dictionary*. 5<sup>th</sup> edition. 1<sup>st</sup> edition 1987. Glasgow: HarperCollins Publishers.

Turnbull, Joanna, ed. 2010. *Oxford Advanced Learner's Dictionary of Current English*. 8<sup>th</sup> edition. 1<sup>st</sup> edition 1948. Oxford: Oxford University Press.

### Schoolbooks

Folland, Tarja, Arja Haavisto, Mari Heinonen, Arto Nieminen and Mark Woods. 2005. *Smart Moves 1. Texts*. Helsinki: Otava.

Folland, Tarja, Arja Haavisto, Mari Heinonen, Arto Nieminen and Mark Woods. 2009. *Smart Moves 1. Exercises*. Fifth edition. Helsinki: Otava.

Folland, Tarja, Arja Haavisto, Tiina Huohvainen, Arto Nieminen, Karolina Sveiby and Marjut Vaakanainen. 2007. *Smart Moves 2. Texts*. Helsinki: Otava.

Folland, Tarja, Arja Haavisto, Tiina Huohvainen, Arto Nieminen and Marjut Vaakanainen. 2007. *Smart Moves 2. Exercises*. Helsinki: Otava.

Folland, Tarja, Tiina Huohvainen, Arto Nieminen, Karolina Sveiby and Marjut Vaakanainen. 2008. *Smart Moves 3. Texts*. Helsinki: Otava.

Folland, Tarja, Tiina Huohvainen, Arto Nieminen and Marjut Vaakanainen. 2008. *Smart Moves 3. Exercises*. Helsinki: Otava.

Haapala, Mika, Raija Kangaspunta, Eero Lehtonen, Jyrki Peuraniemi, Leena Semi and Paul Westlake. 2009. *Spotlight 7: Faces and places. Textbook*. Helsinki: WSOYpro Oy.

Haapala, Mika, Raija Kangaspunta, Eero Lehtonen, Jyrki Peuraniemi, Leena Semi and Paul Westlake. 2010. *Spotlight 8: Sights and sounds. Textbook*. Helsinki: WSOYpro Oy.

Haapala, Mika, Raija Kangaspunta, Eero Lehtonen, Jyrki Peuraniemi, Leena Semi and Paul Westlake. 2012. *Spotlight 7: Faces and places. Workbook*. 4<sup>th</sup> Edition. Helsinki: Sanoma Pro Oy.

Haapala, Mika, Raija Kangaspunta, Eero Lehtonen, Jyrki Peuraniemi, Leena Semi and Paul Westlake. 2012. *Spotlight 8: Sights and sounds. Workbook*. 2<sup>nd</sup> Edition. Helsinki: Sanoma Pro Oy.

Haapala, Mika, Raija Kangaspunta, Eero Lehtonen, Jyrki Peuraniemi, Leena Semi and Paul Westlake. 2012. *Spotlight 9: Fact and fiction. Workbook*. Helsinki: Sanoma Pro Oy.

Westlake, Paul, Arja Haavisto, Raija Kangaspunta, Eero Lehtonen and Jyrki Peuraniemi. 2002. *Key English 7: Courses 1-2 Textbook*. Helsinki: WSOY.

Westlake, Paul, Raija Kangaspunta, Eero Lehtonen and Jyrki Peuraniemi. 2003. *Key English 8: Courses 3-5 Textbook*. Helsinki: WSOY.

Westlake, Paul, Merja Auvinen, Raija Kangaspunta, Eero Lehtonen and Jyrki Peuraniemi. 2004. *Key English 9: Courses 6-8 Textbook*. Helsinki: WSOY.

Westlake, Paul, Arja Haavisto, Raija Kangaspunta, Eero Lehtonen and Jyrki Peuraniemi. 2011. *Key English 7: Courses 1-2 Workbook*. Helsinki: WSOYpro Oy.

Westlake, Paul, Raija Kangaspunta, Eero Lehtonen and Jyrki Peuraniemi. 2012. *Key English 8: Courses 3-5 Workbook*. Helsinki: Sanoma Pro Oy.

Westlake, Paul, Merja Auvinen, Raija Kangaspunta, Eero Lehtonen and Jyrki Peuraniemi. 2012. *Key English 9: Courses 6-8 Workbook*. Helsinki: Sanoma Pro Oy.

### Other Works Cited

Andersson, Lars-Gunnar and Peter Trudgill. 1992. *Bad Language*. London: Penguin Books.

Battistella, Edwin L. 2005. *Bad Language: Are Some Words Better than Others?* New York: Oxford University Press.

Bell, Jan and Roger Gower. 1998. "Writing course materials for the world: a great compromise". In *Materials Development in Language Teaching*, ed. Brian Tomlinson, 116-129. Cambridge: Cambridge University Press.

Biber, Douglas, Susan Conrad and Randi Reppen. 1998. *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press.

Biber, Douglas and Randi Reppen, eds. 2012a. *Corpus Linguistics. Volume I: Lexical Studies*. London: SAGE Publications Ltd.

Biber, Douglas and Randi Reppen, 2012b. "What Does Frequency Have to Do with Grammar Teaching?". In *Corpus Linguistics. Volume IV: Methods and Applications*, ed. Douglas Biber and Randi Reppen, 239-248. London: SAGE Publications Ltd.

Brumfit, Christopher J., Anita Pincas and Geoffrey Broughton. 1978. *Teaching English As a Foreign Language*. London: Routledge.

Carter, Ronald. 1998a. *Vocabulary : Applied Linguistic Perspectives*. London: Routledge.

Carter, Ronald. 1998b. "Orders of reality: CANCODE, communication and culture." *ELT Journal*, 52, 1: 43-56.

Dumas, Bethany K. and Jonathan Lighter 1978. "Is *Slang* a Word for Linguists?". In *American Speech*, 53, 1: 5-17.

Fox, Gwyneth. 1998. "Using corpus data in the classroom". In *Materials Development in Language Teaching*, ed. Brian Tomlinson, 25-43. Cambridge: Cambridge University Press.

- Jackson, Howard. 2002. *Lexicography: An Introduction*. London: Routledge.
- Jackson, Howard and Etienne Zé Amvela. 2000. *Words, Meaning and Vocabulary*. London: Cassell.
- Landau, Sidney I. 1989. *Dictionaries: The Art and Craft of Lexicography*. Cambridge: Cambridge University Press.
- Martin, J.R. 2013. "Embedded literacy: Knowledge as meaning". In *Linguistics and Education*, 24, 1: 23-37.
- McCarten, Jeanne. 2007. *Teaching Vocabulary. Lessons from the Corpus, Lessons for the classroom*. Cambridge: Cambridge University Press.
- Mishan, Freda. 2004. *Designing Authenticity into Language Learning Materials*. Bristol: Intellect Ltd.
- Moilanen, Kari. 2002. *Yli esteiden. Oppimisvaikeudet ja vieraat kielet*. Helsinki: Kustannusosakeyhtiö Tammi.
- Moon, Rosamund. 2007. "Phraseology in general monolingual dictionaries". In *Phraseology: an international handbook of contemporary research. Volume 2*, ed. Harald Burger, 909-918. Berlin: de Gruyter.
- O'Keeffe, Anne and Fiona Farr. 2003. "Using language corpora in initial teacher education: Pedagogic issues and practical applications". *TESOL Quarterly* 37, 3: 389-418.
- Römer, Ute. 2008. "Corpora and language teaching". In *Corpus Linguistics. An International Handbook. Volume 1*, ed. Anke Lüdeling and Merja Kytö, 112-130. Berlin: Mouton de Gruyter.
- Römer, Ute. 2004. "Comparing real and ideal language learner input: The use of an EFL textbook corpus in corpus linguistics and language teaching". In *Corpora and Language Learners*, ed. Guy Aston, Silvia Bernardini, and Dominic Stewart, 151-168. Amsterdam: John Benjamins Publishing Company.
- Thim, Stefan. 2012. *Phrasal Verbs: The English Verb-Particle Construction and Its History*. Berlin: De Gruyter Mouton.
- Wild, Kate. 2011. "Phrasal verbs: 'a process of the common, relatively uneducated, mind'?" In *English Today*, 27, 4: 53-57.

### Electronic Sources

- Davies, Mark. (2008a-) *The Corpus of Contemporary American English: 450 million words, 1990-present*. Available online at <http://corpus.byu.edu/coca/>.
- Davies, Mark. (2008b-) *The Corpus of Contemporary American English: 450 million words, 1990-present*. "Introduction". Available online at <http://corpus.byu.edu/coca/>.
- Davies, Mark. (2008c-) *The Corpus of Contemporary American English: 450 million words, 1990-present*. "Composition of the corpus". Available online at <http://corpus.byu.edu/coca/>.



Davies, Mark. (2008d-) *The Corpus of Contemporary American English: 450 million words, 1990-present*. "Spoken transcripts". Available online at <http://corpus.byu.edu/coca/>.

Davies, Mark. (2008e-) *The Corpus of Contemporary American English: 450 million words, 1990-present*. "Help: Search syntax". Available online at <http://corpus.byu.edu/coca/>.

Davies, Mark. (2010) *Word frequency data - Corpus of Contemporary American English*. Available online at <http://www.wordfrequency.info/free.asp>

Davies, Mark. (2013) *Corpus of Global Web-Based English: 1.9 billion words from speakers in 20 countries*. Available online at <http://corpus.byu.edu/glowbe/>.

EF Education First. 2014a. "EF EPI – English Proficiency Index. Home". Available from <http://www.ef.edu/epi/> [Accessed: 17.04.2015]

EF Education First. 2014b. "EF EPI – English Proficiency Index. About EF EPI: Methodology". Available from <http://www.ef.edu/epi/about-epi/methodology/> [Accessed: 17.04.2015]

Finlex. 1998. "Basic Education Act". Edita Publishing Oy. Available from <http://www.finlex.fi/en/laki/kaannokset/1998/en19980628.pdf> [Accessed: 04.12.2014]

Finnish National Board of Education (Opetushallitus). 2012. "National Core Curriculum for Basic Education 2004". Available from [http://www.oph.fi/download/47672\\_core\\_curricula\\_basic\\_education\\_3.pdf](http://www.oph.fi/download/47672_core_curricula_basic_education_3.pdf) [Accessed: 22.05.2013]

Finnish National Board of Education (Opetushallitus). 2014a. "Basic education". Available from [http://www.oph.fi/english/curricula\\_and\\_qualifications/basic\\_education](http://www.oph.fi/english/curricula_and_qualifications/basic_education) [Accessed: 04.12.2014]

Finnish National Board of Education (Opetushallitus). 2014b. "Education policy". Available from [http://www.oph.fi/english/education\\_system\\_education\\_policy](http://www.oph.fi/english/education_system_education_policy) [Accessed: 04.12.2014]

Finnish National Board of Education (Opetushallitus). 2014c. "The Finnish National Board of Education". Available from [http://www.oph.fi/english/about\\_fnbe](http://www.oph.fi/english/about_fnbe) [Accessed: 04.12.2014]

Opetushallitus (Finnish National Board of Education). 2004. "Perusopetuksen opetussuunnitelman perusteet" ("National Core Curriculum for Basic Education"). Available from [www.oph.fi/download/139848\\_pops\\_web.pdf](http://www.oph.fi/download/139848_pops_web.pdf) [Accessed: 23.09.2012]

Yle Uutiset. 2015. "Finns top-notch in English skills". Available from: [http://yle.fi/uutiset/finns\\_top-notch\\_in\\_english\\_skills/7934886](http://yle.fi/uutiset/finns_top-notch_in_english_skills/7934886) [Accessed: 17.04.2015]

## Appendix I

Word lists used in the study.

### SMART MOVES 1

#### 2. The street where I live

large  
called  
pest  
follow  
most of the time  
guess  
pretty  
size  
tall  
fat  
thin  
wear, wore, worn  
skater  
parents  
businessman  
translator  
building  
suburb  
detached house  
only child  
hang out  
hang, hung, hung  
apartment building  
neighbourhood  
brick  
storey  
bright  
purple  
Portugese  
Italian  
imagine  
fit in  
three doors down  
cook  
meat sauce  
gardener  
grow, grew, grown  
vegetable  
backyard  
cute  
Greek  
twin  
restaurant  
downtown  
street hockey  
actually  
schoolyard  
driveway  
get time of  
join in  
net  
post  
hockey stick  
blade  
break, broke, broken  
replacement  
goalie  
fill up  
goalie  
shoot the ball  
shoot, shot, shot  
beat, beat, beaten  
get into fights  
NHL (National Hockey League)  
clear the snow  
snow bank  
hockey rink  
puck  
freeze, froze, frozen  
hard as rock  
dream, dreamt, dreamt

iso, suuri  
nimeltään  
kiusankappale  
seurata  
enimmäkseen  
arvata ; luulla  
(tässä:) melko  
koko  
pitkä; korkea  
lihava; paksu  
laiha  
pitää yllään; käyttää  
rullaluistelija  
vanhemmat  
liikemies  
kääntäjä  
rakennus  
esikaupunki  
omakotitalo  
ainoa  
olla kavereiden kanssa

kerrostalo  
naapuristo  
tiili  
kerrostalo  
kirkas  
purppuranpunainen  
portugalilainen  
italialainen  
kuvitella  
sopia  
kolmen talon päässä  
kokki  
pastakastike (lihasta)  
puutarhuri

vihannes  
takapiha  
söpö  
kreikkalainen  
kaksonen  
ravintola  
keskusta  
katukiekkko  
itse asiassa  
koulun piha  
pihatie  
saada vapaata  
liittyä seuraan  
verkko  
tolppa  
maila  
lapa  
mennä rikki  
vara(lapa)  
maalivahti  
täyttää  
maalivahti  
ampua pallo

voittaa; lyödä  
joutua tappeluun  
Kanadan ja USA:n jääkiekkoliiga  
aurata  
lumivalli  
jääkiekkokaukalo  
kiekko  
jäättyä  
kivikova  
unelmoida

## SMART MOVES 1

### 12. School the Welsh way

break  
lunch break  
subject  
maths  
geography  
biology  
finish

Meet Sion Davies

Sion Davies  
capital  
Wales  
Ysgol Glantaf  
mean, meant, meant  
secondary school  
to teach  
teach, taught, taught  
Welsh  
speak, spoke, spoken  
most  
come, came, come  
Geraldine  
Sean  
Scottish  
Irish  
surname  
common  
actually  
another  
mate  
Dafydd

Getting to school  
get up  
have to  
wait( for)  
we are in hurry  
arrive  
just in time  
registration with our class teacher  
understand, understood, understood  
it only takes me  
take, took, taken  
from start to finish  
twice as long  
suppose  
spen, spent, spent,  
with her hair and make-up  
luckily  
choose, chosen, chosen  
wear, wore, worn  
school uniform

going to school  
lesson  
most of  
rugby  
guess  
muddy  
trousers  
national sport  
basketball  
cricket  
hockey

to each other

interested (in)  
as well as  
netball

Why study Welsh?

after school  
quite a lot  
go on  
go, went, gone  
outside  
however

välitunti, tauko  
ruokatunti  
(opiskeltava) aine  
matematiikka  
maantieto  
biologia  
loppua

Saanko esitellä Sion  
Daviesin?

*pojan nimi*  
pääkaupunki  
Wales  
*koulun nimi*  
tarkoittaa  
yläkoulu  
jossa opetetaan

walesin kieli; walesilainen

useimmat

*tytön nimi*  
*pojan nimi*  
Skotlannissa puhuttu keltin kieli  
Irlannissa puhuttu keltin kieli, iiri  
sukunimi  
yleinen, tavallinen  
itse asiassa  
toinen  
kaveri  
*pojan nimi*

Miten kouluun mennään  
nousta ylös vuoteesta  
täytyy  
odottaa (jotakin)  
meillä on kiire  
saapua  
juuri ajoissa, viime hetkellä  
ilmottautuminen luokanvalvojalle  
ymmärtää  
minulta kestää vain

alusta loppuun  
kaksi kertaa niin kauan  
olettaa, arvella  
kuluttaa, viettää  
hiustenlaitossa ja meikkaamisessa  
onneksi  
valita  
käyttää, pitää yllään  
koulupuku

Koulunkäynti  
oppitunti  
suurin osa  
rugby-jalkapallo  
arvata  
kurainen  
pitkät housut  
kansallisurheilu  
koripallo  
kriketti  
maahockey  
toistensa kanssa,  
keskenään  
kiinnostunut( jostakin)  
kuin myös  
tyttöjen koripallo

Miksi walesin kieltä  
kannattaa opiskella?  
(tässä:) lopetettuaan  
koulun  
aika monet  
jatkaa eteenpäin

ulkopuolella,  
ulkopuolelle  
siitä huolimatta,  
kuitenkaan

a waste of time  
anyway  
know, knew, known  
world  
through  
have a point

School uniform  
navy  
skirt  
sensible  
length  
classic style  
pale blue  
shirt  
badge  
neck  
sweatshirt  
fleece  
anorak  
trainers  
high heels  
sock  
hairstyle  
acceptable  
jewellery  
be allowed  
apart from  
plain stud earrings

ajanhukkaa, tarpeetonta  
joka tapauksessa  
(tässä:) osata  
maailma  
läpi  
olla oikeassa, puhua järkeä

Koulupuku  
tummansininen  
hame  
järkevä, asiallinen  
pituus  
tavalliset suorat pitkät housut  
hennon vaalean siniseen  
kauluspaita  
tunnusmerkki, vaakuna  
(tässä:) kaula-aukko  
college-pusero  
fleece-takki  
tuulen ja sateen kestävä takki  
lenkkarit  
korkokengät  
sukka  
hiusmalli, kampa  
hyväksyttävä, sopiva  
korut  
sallia  
paitsi  
korvanappikorut

## SMART MOVES 2

### 17. Land of shamrocks and leprechauns -Excursion to Ireland

Excursion to Ireland  
diary

bring, brought, brought  
classmate  
excited (about)  
crack a joke  
airport  
through  
passport control  
security check  
departure gate  
boarding card  
get on the plane  
fly, flew, flown  
change planes  
airline  
flight attendant  
make jokes

land  
look around  
hardly  
double-decker  
watch out (for)  
on the left-hand side  
coach ride  
hometown  
exhausted  
since one o'clock  
most of the way  
miss  
scenery  
understand, understood, understood  
give a talk  
point (at)  
on the way  
arrive (at)  
hostel  
climb  
gear  
sight  
bunk bed  
bleak  
luxury  
still  
a toilet of our town  
work

Leirikoulu Irlantiin  
päiväkirja

tuoda  
luokkakaveri  
innostunut, jännittynyt  
(jostakin)  
kertoa, 'murjaista', vitsi  
lentokenttä  
läpi  
passitarkastus  
turvatarkastus  
lähtöportti  
matkustajan  
tarkistuskortti  
mennä lentokoneeseen  
lentää  
vaihtaa konetta  
lentoyhtiö  
lentöemäntä tai stuertti  
laskea leikkiä, vitsailla

laskeutua  
(tässä:) katsella  
nähtävyyksiä  
tuskin  
kaksikerroksinen bussi  
varoa (jotakin)  
vasemmalla  
bussimatka  
kotikaupunki  
loppuunpunut  
kello yhdestä lähtien  
suurimman osan matkasta  
mennä joltakin ohi  
maisema  
ymmärtää  
pitää esitelmä  
osoittaa (jotakin)  
matkalla  
saapua (jonnekin)  
nuorisohotelli  
kiivetä  
varusteet, tavarat  
näky, näkeminen  
kerrossänky  
kolkko  
yllellisyys  
silti, kuitenkin  
oma vessa  
(tässä:) toimia

Whatever the beds were like  
knackered  
travel  
sleep like a log

highlight  
start  
sightseeing  
bay  
cave  
cliff  
unusual  
grave  
nothing but  
footprint  
underground  
limestone  
windy  
abbey  
nun  
stormy  
boat ride  
foggy  
fort  
quicksand  
freezing  
farewell dinner  
gift  
coach  
canteen

## SMART MOVES 2

### 1. Aussie way of life

opinion  
wildlife  
Down Under  
exotic  
agree  
disgusting  
though  
exciting  
cute  
scary  
give up  
We agree to disagree!

exchange student  
spend, spent, spent  
fact  
stay with  
host brother

notice  
Australian  
mad  
I guess  
almost  
do sport  
do, did, done  
popular  
follow  
footy  
cricket  
favourite  
as well  
I'm afraid  
rule  
at all  
be into  
that's more like it  
hear, heard, heard  
crazy  
event  
race  
regatta  
completely  
cockroach  
actually  
hold, held, held  
become, became, become  
coach

Millaisia tahansa sängyt olivatkaan  
poikki, 'rättiväsynyt'  
matkustaa  
nukkua kuin tukki

kohokohta  
aloitus  
nähtävyyksien katseleminen  
lahti  
luola  
kallioseinä, rantakallio  
epätavallinen  
hauta  
ei muuta kuin  
jalanjälki  
maan alla  
kalkkikivi  
tuulinen  
luostari  
nunna  
myrskyinen  
venematka  
sumuinen  
linnoitus, linnake  
juoksuhiikka  
jäätyä kylmä  
lähtöillallinen  
lahja  
(bussi) kuljetus  
ruokailuhuone, ruokala

mielipide  
villieläimet ja kasvit  
Australia (ja Uusi-Seelanti)  
eksoottinen, erikoinen  
olla samaa mieltä  
inhottava  
kuitenkin  
jännittävä  
söpö  
pelottava  
luovuttaa  
Olemme yhtä mieltä siitä,  
että olemme eri mieltä asiasta!

vaihto-oppilas  
viettää (aikaa)  
tosiasia, fakta  
asua (tilapäisesti) jonkun luona  
isäntäperheen poika

huomata  
australialainen  
hullu  
(tässä:) varmaankin  
melkein  
harrastaa urheilua  
tehdä  
suositettu  
seurata  
australialainen jalkapallo  
kriketti  
suositettu, lempi-  
myös  
valitettavasti  
sääntö  
lainkaan, ollenkaan  
harrastaa, olla innostunut (jostakin)  
se kuulostaa jo paremmalta  
kuulla  
hullu  
tapahtuma; laji  
kilpailu ; kilpailu  
regatta, purjehdus-/  
soutukilpailut  
täysin  
torakka  
itse asiassa, todellakin  
(tässä:) järjestää  
tulla joksikin  
valmentaja

Our globetrotter in Australia

article  
magazine  
experience  
G'day  
the outback  
change of landscape  
capital  
adventure  
mostly  
which  
outdoor activities  
hot air balloon  
ride, rode, ridden  
quad bike  
hire  
four-wheel-drive  
bush  
degree  
trust me  
sigh  
local  
guide  
Aboriginal  
background  
take, took, taken  
Uluru  
holy  
rock  
native Australian  
respect  
climb  
on your own  
poisonous  
camp  
the Southern Cross  
wherever  
carefree  
sense of humour  
feel at home  
Take care!

aboriginal significance

Captions  
bottomless  
race  
hold on (to)

Maailmanmatkaajamme

Australiassa  
artikkeli, kirjoitelma  
aikakauslehti  
kokea; kokemus  
hei, tervehdys (Austr.)  
Australian erämaa  
maiseman muutos  
pääkaupunki  
seikkailu  
enimmäkseen  
joka  
ulkoilmaharrastukset  
kuumailmapallo  
ajaa (pyörällä); ratsastaa  
mönkijä  
vuokrata  
nelivetoinen auto  
(harvaanasuttu) metsikkö  
(Austr.)  
aste  
usko pois  
nähtävyys  
paikallinen  
opas  
alkuasukas- (Austr.)  
tausta  
viedä; ottaa  
Australian alkuasukkaiden  
pyhä kallio  
pyhä  
kallio, (iso) kivi  
Australian alkuperäiskansa  
kunnioittaa  
kiivetä  
yksinään, omin päin  
myrkyllinen  
leiriytyä  
Etelän risti (tähtikuvio)  
missä tahansa; minne tahansa  
rento, huoleton  
huumorintaju  
tuntea olonsa kotoisaksi  
Pärjää!

alkuasukille tärkeä

Kuvatestit  
pohjaton  
kilpailu  
pitää kiinni (jostakin)

## SMART MOVES 3

### 23. The internet - sex, lies, and deceit

everyday  
tool  
communication  
according to  
recent study  
percentage  
channel  
vital  
source  
highway  
playground  
fraud  
wacko  
tackle  
authorities  
set up  
promote  
usage  
target  
download  
update  
operating system  
virus protection  
up-to-date  
firewall

päivittäinen  
työkalu  
viestintä  
jonkun mukaan  
viimeaikainen tutkimus  
prosenttiosuus  
kanava  
elintärkeä  
lähde  
valtatie  
leikkikenttä  
huijari; petos  
hullu  
ottaa käsittelyä, hoitaa  
viranomaiset  
pystyttää; ottaa käyttöön  
edistää, tukea  
käyttö  
tavoitella  
suunnata toiminta  
ladata  
päivitys  
käyttöjärjestelmä  
virustorjunta  
ajan tasalla  
palomuuuri

block	estää pääsy
intruder	tunkeilija
be aware (of)	olla tietoinen (jostakin)
include	pitää sisällään
spamming	roskapostin lähettäminen
unwanted	ei-toivottu
spamming	roskaposti
inbox	saapuneet-postilaatikko
estimate	arvioida
fourth	neljäsosa
advertise	mainostaa
law service	lakipalvelu
low cost	halpa
advertising	mainonta
equally	yhtä
potential	mahdollinen
single	yksittäinen
ad=advert=advertisement	mainos
pointless	järjetön
annoying	ärsyttävä, kiusallinen
so-called	niin kutsuttu
discreetly	hienovaraisesti
account	tili
sum	summa
fall for the scam	mennä lankaan
document	asiakirja
proof	todiste
exist	olla olemassa
delay	viivytyt
in order to	jonkun asian toteutumiseksi
bribe	lahjoa
official	viranomainen
lucky break	onnenpotku
invest (in)	sijoittaa (johonkin)
deal-of-a-lifetime	elämän paras sijoitus/kauppa
opportunity	mahdollisuus
your way	sinun suuntaasi
Phishing	kalastelu
criminal activity	(tietotekniikassa)
masquerade (as)	rikollistoiminta
representative	esiintyä (jonakin)
trustworthy	edustaja
false identity	luotettava
acquire	väärä henkilöllisyys
credit card	saada haltuunsa
password	luottokortti
client	tunnussana
log on	asiakas
given	kirjautua sisään
personal details	annettu
via	henkilötiedot
	(jonkun) kautta
Search engine use	Hakukoneiden käyttö
simply	yksinkertaisesti
browse	selailla
catalogue	luetteloida
search box	hakulaatikko
compare	verrata
hit	osuma
porn	porno
achieve	saavuttaa
add	lisätä
Parody pages	parodiasivut
encyclopaedia	tietosanakirja
edit	muokata
be about	olla kyseessä
resource	lähde
come across	kohdata sattumalta
content-free	sisällötön
False information	väärää tietoa
contain	sisältää
half-truth	puolitotuus
lie	valhe
censor	sensuroida
therefore	tämän vuoksi
require	vaatia
pay attention (to)	kiinnittää huomio (johonkin)
content	sisältö
scepticism	skeptisismi, epäily
urban myth	kaupunkilaistarina

copy and paste  
endless  
thus  
create  
illusion  
portray  
integral  
harm

leikata ja liimata  
loputon  
täten  
luoda  
harhakuvitelma  
esittää  
olennainen, erottamaton  
satuttaa

### SMART MOVES 3

#### 4. Global issues: energy

energy  
greenhouse effect  
global warming

energia  
kasvihuoneilmiö  
maapallon lämpeneminen

Sources of energy  
fossil fuel  
coal  
oil  
natural gas  
peat  
limited amount  
uranium  
nuclear energy  
hydro power  
solar power  
again and again  
form  
biomass energy  
geothermal energy  
one fifth  
electricity  
in the future  
another  
windmill park  
refer (to)  
wood  
manure  
to a great extent  
basically  
earth-heat  
the further - the warmer  
underground

Energianlähteitä  
fossiilinen polttoaine  
kivihiili  
öljy  
maakaasu  
turve  
rajallinen määrä  
uraani  
ydinenergia  
vesivoima  
aurinkovoima  
yhä uudelleen  
muoto  
bioenergia  
maalämpöenergia  
yksi viidesosa  
sähkö  
tulevaisuudessa  
(vielä) yksi  
tuulimyllypuisto,  
tuulivoimala  
viitata (johonkin)  
puu  
lanta  
suuressa määrin  
pohjimmiltaan  
maalämpö  
mitä edemmäs - sitä lämpimämpää  
maan alla, maan alle

The greenhouse effect  
planet  
work  
warming radiation  
enter  
breathe  
warmth  
survive  
gas  
prevent (from)  
escape  
in other words  
freezing

Kasvihuoneilmiö  
planeetta  
toimia  
lämpösäteily  
tulla (sisään)  
hengittää  
lämpösäteily  
selviytyä  
kaasu  
estää (jostakin)  
paeta, karata  
toisin sanoen  
jäättävä

Global warming  
extra  
as a result  
temperature  
rise, rose, risen  
risk  
glacier  
melt  
melting process  
so far  
consequence  
dramatic  
continue  
sea level  
cause damage  
lowland area  
Southeast England  
increase  
consider

Maapallon lämpeneminen  
ylimääräinen  
seurauksena  
lämpötila  
nousta  
riski, uhka  
jäättikkö  
sulaa  
sulamisprosessi  
tähän mennessä, toistaiseksi  
seuraus  
dramaattinen, todella suuri  
jatkua  
merenpinnan taso  
aiheuttaa vahinkoa  
alava maa-alue  
Kaakkois-Englanti  
lisätä  
harkita

Pollution  
harmful  
particle  
float

Saastuminen  
haitallinen  
hiukkanen  
leijua



asthma  
get worse  
acid rain  
falling  
poisonous  
in addition  
rivers and lakes with acidity levels

astma  
pahentua  
happosade  
laskeutuva  
myrkyllinen  
lisäksi  
happamat joet ja järvet  
(pH alle 7)

## KEY ENGLISH 1 1A Kids on the move

a kid  
a move  
on the move  
meet, met  
different  
but  
a thing  
common  
have in common  
thirteen  
go, went through  
a change  
a life  
a city  
a city slicker  
a bumpkin  
England  
a location  
a village  
called  
a comprehensive (school)  
large  
a detached house  
dad  
Tottenham Hotspurs  
dislike  
Arsenal  
people  
move  
London  
last  
a summer  
last summer  
believe  
very  
here  
Cornwall  
have, had to  
get, got up  
early  
a bus  
catch the school bus  
everybody  
laugh at  
a way  
speak, spoke  
the ninth  
a floor  
twenty  
a storey  
Twenty-storey  
a towerblock  
near  
a football ground  
Eight-roomed  
in the middle of  
nowhere  
in the middle of nowhere  
more  
a lot more room  
giant  
a garden  
much  
hard  
harder  
work  
keep, kept  
a place  
clean  
tidy

lapsi, nuori  
muutto, siirto  
liikkeellä, liikekannalla  
tavata, kohdata  
erilainen, eri  
mutta  
asia, esine  
yhteinen, tavallinen  
olla yhteistä  
kolmetoista  
käydä läpi, kokea  
muutos  
elämä  
kaupunki  
kaupunkilainen, jupi  
maalainen, juntti  
Englanti  
sijainti  
kylä  
nimeltä, jonka nimi on  
(Br.) Koulu 11-18 – vuotiaille, peruskoulu  
suuri, iso  
omakotitalo  
isä  
engl.jalkapallojoukkue  
inhota, ei pitää jostakin  
engl.jalkapallojoukkue  
ihmiset  
muuttaa  
Lontoo  
viime  
kesä  
viime kesänä  
uskoa  
hyvin, kovin  
täällä, tässä, tähän  
  
täytyä  
nousta ylös  
aikainen, aikaisin  
linja-auto, bussi  
ehtiä koulubussiin  
jokainen, kaikki  
nauraa jollekin  
tapa  
puhua  
yhdeksäs  
kerrros, lattia  
kaksikymmentä  
kerros  
kaksikymmenkerroksinen  
kerrostalo  
lähellä  
jalkapallokenttä  
kahdeksan huoneen  
jonkin keskellä  
ei missään  
syrjäisessä paikassa,  
kaukana kaikesta  
enemmän  
paljon enemmän tilaa  
jättiläismäinen, valtavan iso  
puutarha, piha  
paljon  
kova, ankara  
kovempi  
työ  
pitää  
paikka  
puhdas, siisti  
siisti

Mind you!  
a room  
that  
say, said  
that  
new  
more than  
a primary school  
tease  
all the time  
now  
small  
smaller  
only  
a pupil  
a class  
not only that but  
Well-organized  
brand new  
what  
a difference  
What a difference!  
so  
on the whole  
happy  
happier  
feel, felt  
healthy  
healthier  
too  
Canada  
an apartment  
mom  
hockey  
canoeing  
traffic  
violence  
TV  
a soap opera  
high school  
just  
outside  
seem like  
another  
a world  
a bit  
difficult  
adjust to  
a neighbourhood  
busy  
busier  
Things are busier here in Edmonton  
people  
less  
friendly  
we used to live  
all  
there  
share a room  
with  
young  
younger  
tell, told  
keep telling  
a door  
closed  
no one  
try, tried  
no one has tried  
bully  
at school  
probably  
because  
build, built  
be built like a tree  
some  
a guy  
a hick  
button  
a lip  
button his lip  
buzz off

Ota huomioon! Muuten...  
huone  
tuo, se  
sanoa  
että  
uusi  
enemmän kuin  
(Br.) Koulu 5-11 – vuotiaille, alakoulu  
(Suomi)  
kiusata, kiusoitella  
koko ajan  
nyt  
pieni  
pienempi  
vain  
oppilas  
luokka  
sen lisäksi  
hyvin järjestetty, järjestelmällinen  
ihan uusi  
mikä, mitä  
ero  
Mikä ero! Valtava ero!  
joten; siis; niin  
kaiken kaikkiaan  
onnellinen  
onnellisempi  
tuntua, tuntea olonsa joksikin  
terve, terveellinen  
terveempi, terveellisempi  
myös,-kin  
Kanada  
huoneisto, asunto  
(Am.) äiti  
(Am.) jääkiekko  
melonta  
liikenne  
väkivalta  
TV, televisio  
saippuasarja, saippuaoppera  
(Am) koulu 14–18-vuotiaille  
juuri, aivan, vain  
ulkopuolella, ulkopuolelle  
näyttää joltakin,  
tuntua joltakin  
toinen  
maailma  
hieman  
vaikea  
sopeutua johonkin  
asuinalua  
kiireinen  
kiireisempi  
Täällä Edmontonissa on  
kiireisempää.  
ihmiset  
vähemmän  
ystävällinen  
asuumme ennen  
kaikki  
siellä, tuolla  
asua samassa huoneessa  
kanssa  
nuori  
nuorempi  
kertoa, sanoa, käskä  
käskää jatkuvasti  
ovi  
suljettu, kiinni  
ei kukaan  
yrittää, koettaa, kokeilla  
kukaan ei ole yrittänyt  
kiusata  
koulussa  
luultavasti  
koska  
rakentaa  
olla isokokoinen  
joku, jotkut  
tyyppi, jätkä  
juntti  
napittaa, panna napit kiinni  
huuli  
pitää suunsa kiinni  
painua tiehensä

really  
miss  
hope  
can  
pull through  
OK  
shall  
wait  
see, saw  
wait and see  
guess

todella  
kaivata, ikävöidä, menettää  
toivoa  
pystyä, osata, voida  
selviytyä  
hyvin  
pitää  
odottaa  
nähdä  
odottaa mitä tapahtuu  
arvata, arvella

## KEY ENGLISH 1

### Unit 3 School is cool!

cool  
back  
in and out  
a toilet  
up and down  
stairs  
an assembly  
queue  
a pair  
in pairs  
a crush  
cute  
a smile  
gym(nastics)  
I'd run  
a mile  
lunch  
a packed lunch  
a biscuit  
crisps  
an apple  
munch  
munches  
listen to  
a subject  
pretty  
bad  
history  
chemistry  
a fool  
only  
poetry  
a meal  
a pupil  
teach, taught, taught  
a timetable  
a lesson  
an exam  
a test  
homework  
stand, stood, stood  
by  
a desk  
a head  
monday  
tuesday  
wednesday  
thursday  
friday  
a refistration  
art  
a break  
science  
lunch  
history  
german  
drama  
maths=mathematics  
geography  
french  
biology  
P.E.= physical education  
IT=information technology  
music  
chemistry  
enjoy  
attend  
secondary school

viileä, upea, hieno  
takaisin  
sisään ja ulos  
WC  
ylös ja alas  
portaat  
yhteinen päivänavaus,  
kokoontuminen  
jonottaa  
pari  
pareittain  
ihastus  
söö, suloinen  
hymy  
voimistelu  
juoksisin  
maili  
lounas  
eväät, eväspaketti  
keksi  
perunalastut  
omena  
rouskuttaa  
(tässä:) rouskutus  
kuunnella jotakin  
oppiaine  
aika, melko  
huono  
historia  
kemia  
typerys  
ainoa  
runous  
ateria  
oppilas  
opettaa  
lukujärjestys  
oppitunti  
koe, tentti  
koe, tentti  
kotitehtävä, läksy  
seisoa  
luona, vieressä  
pulpetti; kirjoituspöytä  
(br) rehtori  
maanantai  
tiistai  
keskiviikko  
torstai  
perjantai  
ilmottautuinen, nimenhuuto  
kuvataide, taide  
tauko, välitunti  
luonnontieteet  
lounas  
historia  
saksan kieli  
ilmaisutaito  
matematiikka  
maantieto  
ranskan kieli  
biologia  
liikunta  
tietotekniikka  
musiikki  
kemia  
nauttia, pitää jostakin  
osallistua, käydä (koulua)  
(br) koulu 11–16/18

over	-vuotiaille,
generally	yläkoulu (suomi)
a quarter	yli
	yleensä
at a quarter to nine	neljännes, neljännestunti,
mean, meant, meant	vartti
a for teacher	viittätoista vaille yhdeksän,
check	neljännes vaille yhdeksän
if	tarkoittaa
a form	(br.) luokanvalvoja, -ohjaaja
take place	tarkastaa
a form room	jos, -ko, -kö
a hall	(br.) luokka, luokka-aste
a kind of	tapahtua
a get-together	luokkahuone
sing, sang, sung	aula, sali
a song	eräänlainen
read, read, read out	kokoontuminen
a story	laulaa
twice	laulu
during	lukea ääneen
past	kertomus
a month	kaksi kertaa
go off	aikana
different	viime, kulunut
a class	kuukausi
a minute	mennä, lähteä
an afternoon	eri, erilainen
a break	luokka, oppitunti
between	minuutti
finish	iltapäivä
home	tauko, välitunti
a head teacher	välillä, välissä
funny	päätyä, loppua
care about	kotiin
a reason	rehtori
Extra-curricular	hauska
an activity	välittää jostakin/ jostakusta
Extra-curricular activities	syy
	oppituntien ulkopuolinen
a club	toiminta
media studies	koulun harrastus- ja kerhotoiminta
food	(oppituntien ulkopuolella)
nutrition	kerho
unfortunately	viestintä
a period	ruoka
a term	ravinto
enough	valitettavasti
sports	jakso
matter	lukukausi
belong to	tarpeeksi
possible	urheilu
as soon as possible	olla merkitystä, olla väliä
luv=love	kuulua johonkin
a film	mahdollinen
be like	mahdollisimman pian
what time	terveisin
a student	elokuva
a teacher	olla jonkin kaltainen
a month	mihin aikaan
January	opiskelija
February	opettaja
March	kuukausi
April	tammikuu
May	helmikuu
June	maaliskuu
July	huhtikuu
August	toukokuu
September	kesäkuu
October	heinäkuu
November	elokuu
December	syyskuu
short	lokakuu
Christmas	marraskuu
begin, began, begun	joulukuu
hot	lyhyt
between	joulukuu
an autumn	alottaa, alkaa
a holiday	kuuma
after	välillä, välissä
college	(br.) syksy
	loma
	jälkeen
	(Uudessa-seelannissa)
	koulu 14–18 -vuotiaille

Auckland  
 Japanese  
 New Zealand  
 guess  
 a half  
 a year and a half  
 finish  
 in time for  
 a holiday  
 do well  
 most (of)  
 a core subject  
 IT=information technology  
 social science  
 simply  
 a neighbour  
 spend, spent, spent  
 at least  
 talk to  
 easy  
 wear, wore, worn  
 a uniform  
 by the way  
 hate  
 get off  
 get home  
 detention  
 twice  
 once  
 misbehaviour  
 play truant  
 still  
 behaviour  
 last  
 a quarter  
 an hour  
 three quarters of an hour  
 long  
 a rule  
 strict  
 a key word  
 respect (for)  
 others  
 active  
 learning  
 have fun  
 soon  
 see yah=see you  
 a school  
 a schoolbag  
 a book  
 a notebook  
 a file  
 a calendar  
 a pencil case  
 a pencil  
 a pen  
 a felt-tip pen  
 a pencil sharpener  
 a rubber  
 an eraser  
 a ruler  
 a calculator  
 a desk  
 a chair  
 a school yard  
 a classroom  
 a language lab  
 a gym  
 a science lab  
 a computer lab  
 a music room  
 an art room  
 a staff room  
 home economics  
 a library  
 a school canteen  
 a corridos  
 a locker  
 cloakroom(s)  
 a toilet  
 straight  
 line up

japanin kieli; japanilainen  
 Uusi-Seelanti  
 arvata  
 puolikas  
 puolitoista vuotta  
 lopettaa  
 juuri enne, johonkin mennessä  
 loma  
 menestyä  
 useimmat, suurin osa  
 perusoppiaine, kaikille yhteinen  
 oppiaine  
 tietotekniikka  
 yhteiskuntaoppi  
 yksinkertaisesti, vain, pelkästään  
 naapuri  
 kuluttaa, viettää  
 ainakin, vähintään  
 puhua jollekulle  
 helppo  
 käyttää, pitää yllä  
 koulupuku, univormu  
 muuten  
 vihata, inhota  
 ottaa pois, riisua  
 päästä kotiin  
 Jälki-istunto  
 kahdesti  
 kerran  
 huono käytös  
 lintsata koulusta  
 kuitenkin, silti  
 käytös  
 kestää  
 neljännes, vartti  
 tunti  
 neljäkymmentäviisi minuuttia  
 kauan  
 sääntö  
 ankara  
 avainsana, motto  
 kunnioitus (jotakin/  
 jotakuta kohtaan)  
 toiset, muut  
 aktiivinen  
 oppiminen  
 pitää hauskaa  
 pian  
 hei hei, moi  
 koulu  
 koululaukku  
 kirja  
 vihko, muistikirja  
 mappi  
 kalenteri  
 penaali  
 lyijykynä  
 mustekynä  
 huopakynä, tussi  
 kynänteroitin  
 pyyhekumi  
 pyyhekumi  
 viivoitin  
 laskin  
 pulpetti, kirjoituspöytä  
 tuoli  
 koulun piha  
 luokkahuone  
 kielistudio  
 voimistelusal  
 kemianluokka, laboratorioluokka  
 tietokonehuone  
 musiikkiluokka  
 kuvaamataidon luokka  
 opettajainhuone  
 kotitalous  
 kirjasto  
 kouluruokala  
 käytävä  
 lokero, pieni kaappi  
 vaatesäiliö, vaatenaulakot  
 vessa  
 suora, suoraan  
 tehdä jono

sensibly  
cussing  
remove  
hate  
a coat  
interrupt  
without  
follow  
an instruction  
immediately  
hands up

järkevästi  
kiroilu  
ottaa pois  
hattu  
takki  
keskeyttää  
ilman  
seurata  
ohje  
välittömästi  
kädet ylös

## KEY ENGLISH 2 10A Departures

a departure  
a walking holiday  
Denmark  
a tour  
a programme  
travel  
abroad  
live  
Heathrow  
an airport  
  
a transit hall  
  
a passenger  
a traveller  
first  
whose  
a plan  
focus on  
so  
on business  
stay with/in  
of course  
should  
and all that  
main  
look forward to  
excellent  
a river  
a salmon, salmon  
Lapland  
luggage  
a backpack  
full  
check in  
a suitcase  
hopefully  
wait for  
quite a lot  
a Finn  
well enough  
advice  
anyone who  
go, went, gone on holiday  
arrive at  
an hour  
surprise  
a tourist  
a mosquito, mosquitoes  
plenty of  
an insect  
insect repellent  
change  
money  
currency  
an xchange  
the currency exchange  
through  
Customs  
not...yet  
collect  
baggage  
the baggage reclaim  
had better (do)  
let, let, let

lähtö  
patikointiloma  
Tanska  
matka, kiertomatka, kiertue  
ohjelma  
matkustaa, matkustella  
ulkomailla, ulkomaille  
livenä  
  
lentokenttä  
transithalli, lentoaseman  
kauttakulku- ja odotushalli  
jatkolennoilla matkustaville  
matkustaja  
matkustaja, matkailija  
ensiksi, ensin  
jotka, joiden; kenen, keiden  
suunnitelma  
kohdistua johonkin,  
keskittyä johonkin  
niinpä, siis; niin  
liikeasioissa, liikematkalla  
asua (tilapäisesti) jonkun luona/  
hotellissa  
tietenkin  
pitäisi  
ja sen sellaista  
tärkein, pää-  
odottaa kovasti jotakin,  
tuskin maltaa odottaa jotakin  
erinomainen, loistava  
joki  
lohi  
Lappi  
matkatavarat  
rinkka, reppu  
täynnä, täysi  
luovuttaa (matkatavarat)  
lähtöselvityksessä  
matkalaukku  
toivottavasti  
odottaa (jotakin)  
aika paljon, melkoisesti  
suomalainen  
ihan hyvin  
neuvot  
se joka, jokainen joka  
ennä lomalle  
saapua, tulla perille jonnekin  
tunti  
yllättää  
turisti, matkailija  
hyttynen, sääski, moskiitto  
paljon, runsaasti  
hyönteinen  
hyönteiskarkote  
vaihtaa  
raha, rahat  
valuutta  
vaihto, vaihtaminen  
valuutanvaihtopiste  
läpi, kautta  
tulli  
ei vielä  
noutaa, kerätä  
matkatavarat  
matkatavaroiden noutopiste  
on / olisi parasta (tehdä)  
antaa, päästää

forget, forgot, forgotten  
show, showed, shown  
a passport  
Copenhagen  
a resort  
a granny  
a grandmother  
look forward to  
the Canary Islands  
a bit  
be in a hurry  
get through  
a departure lounge  
a gate  
a flight  
have a look at  
a screen  
board  
boarding  
get a move on  
a boarding pass  
a landing  
Happy landings!  
back  
a studio

unohtaa  
näyttää, esittää  
passi  
Kööpenhamiina  
lomakohde  
mummu  
isoäiti  
näyttää  
Kanariansaaret  
vähän, hieman  
olla kiire  
päästä läpi, ehtiä  
lähtöaula  
portti  
lento  
katsoa jostakin  
kuvaruutu, näyttö, monitori  
mennä lentokoneeseen/junaan/laivaan  
siirtyminen lentokoneeseen  
pitää kiirettä  
(lentomatkustajien) tarkistuskortti  
laskeutuminen  
Hyvää matkaa!  
takaisin  
[no translation  
provided]

## KEY ENGLISH 2

### 13B A whole new ball game

whole  
whole new  
a ball game  
a whole new ball game  
take a look at  
major  
invent  
baseball  
basketball

national  
a pastime  
history  
lead, led, led  
a Yankee  
New York Yankees  
a series, series  
a title  
a World Series title

amazing  
be involved  
a hero  
because of  
incredible  
a home run  
a record  
in case  
a batter  
hit, hit, hit  
a field  
out of the field  
a dozen  
dozens of  
professional  
a league  
join  
a little league (AmE)  
look forward to  
Wednesday

find out  
a story  
the story goes that  
way back in 1891  
by the name of  
physical education  
a lesson  
rowdy  
a class  
local  
the YMCA

kokonainen, koko; aivan  
aivan uusi  
pallopeli  
(myös:) aivan uusi asia  
tarkastella jotakin, luoda  
katsaus johonkin  
tärkein, suuri  
keksiä  
baseball  
koripallo

kansallinen, kansallis-  
harrastus  
historia  
johtaa, johdattaa  
jenkki  
amerikkalainen baseball-seura  
sarja  
otsikko, titteli, mestaruus  
(baseballissa)  
maailmanmestaruus  
yllättävä, hämästyttävä  
olla mukana, osallistua  
sankari  
jonkin johdosta, jonkin vuoksi  
uskomaton, aivan mahtava  
(baseballissa) kunniajuoksu  
ennätys  
mikäli, siltä varalta että  
(palloleleissä) lyöjä  
lyödä  
kenttä  
pelialueen ulkopuolelle  
tusina  
kymmeniä, suuria määriä  
ammattilainen  
liiga, sarja  
liittyä johonkin  
(baseballissa) lasten sarja  
odottaa innosta jotakin  
keskiviikko

ottaa selvää, saada selville  
tarina, kertomus, juttu  
tarina kertoo että,  
kerrotaan että  
jo vuonna 1891  
nimeltä  
liikunta (oppiaine)  
(oppi)tunti  
mellastava, villi  
luokka, ryhmä  
paikallinen  
NMKY (=Nuorten miesten kristillinen yhdistys)

control  
a matter  
and to make matters worse  
could  
not even  
outside  
hard  
rugby  
a player  
toss  
a box  
a gymnasium  
nail  
a basket  
high up  
a wall  
throw, threw, thrown  
a team  
make a basket  
a ladder  
climb  
finally  
bright  
cut, cut, cut  
a bottom  
I wonder if  
a millionaire  
ever

still  
a bit of  
a mystery  
explain  
a rule  
as far as  
make out  
a try  
forward  
a yard  
succeed  
keep, kept, kept  
fail  
pass  
eleven  
each  
a quarterback  
a guy  
whether  
rough  
wear, wore, worn  
padding  
the final game  
a bowl

the Super Bowl  
hundreds of millions of  
a television viewer  
be able to  
back home

believe  
bowling  
probably  
popular  
a participant  
a participant sport  
the States  
a spectator  
a spectator sport  
auto-racing  
soccer  
gradually  
overtake, overtook, overtaken  
honest  
to be honest  
number one

### KEY ENGLISH 3

#### 26A The pros and cons of media

pros  
cons  
media  
Good for you

valvoa  
asia  
ja mikä vielä pahempaa  
osasi, voi  
ei edes  
ulos  
kovasti

pelaaja  
heittää, heitellä  
laatikko  
voimistelusalali  
naulata  
kori  
korkealle  
seinä  
heittää  
joukkue  
tehdä kori  
tikapuut  
kiivetä  
lopulta, lopuksi  
loistava  
leikata  
pohja  
(tässä:) ovatkohan  
miljonääri  
koskaan

yhä, vielä  
melkoinen, aikamoinen  
mysteeri, arvoitus, salaisuus  
selittää  
sääntö  
mikäli  
saada selvää, ymmärtää  
yritys  
eteenpäin  
jaardi, n. 90 cm  
onnistua  
pitää  
epäonnistua  
(urheiuksessa) syöttää  
yksitoista  
kukin, kumpikin, joka(inen)  
quarterback, pelinrakentaja  
mies, kaveri, tyyppi  
-kö, -ko  
kova, raju  
olla yllään, käyttää  
toppaukset  
loppuottelu  
malja  
Am. Jalkapallo  
ammattilaissarjan loppuottelu  
satoja miljoonia  
television katsoja  
pystyä, kyetä, voida  
kotipuolella, kotona

uskoa  
keilailu  
luultavasti  
suosittu  
osanottaja, osallistuja  
aktiiviurheilulaji  
Yhdysvallat  
katsoja  
penkkiurheilulaji  
autokilpailut  
jalkapallo  
vähitellen  
ohittaa, voittaa (suosiossa)  
rehellinen  
totta puhuakseni, rehellisesti sanoen  
(tässä:) suosituin

hyvät puolet  
huonot puolet  
tiedotusvälineet



wonderful  
an invention  
entertain  
a companion  
an event  
cheer up  
feel down  
unlike  
switch off  
switch on  
whenever  
feel like it  
a channel  
choose, chose, chosen  
a documentary  
the most  
long live..!

a TV programme  
rubbish  
spend, spent, spent  
horror  
violence  
nearly  
brain  
at  
imagination  
communicate  
naturally  
a human being  
a visitor  
They have been numbed by television  
the box

tune in/into  
listen to  
particularly  
favourite  
news  
weather  
current affairs programme  
after all  
complete  
an airhead  
a board game  
a pal  
a kid sister  
turn off  
rarely  
commercial  
anymore  
junk  
either-or  
a half-brain  
laugh at  
corny  
a joke  
a jerk  
so-called  
drive round the bend  
feed, fed, fed  
read, read, read  
appreciate  
i do read  
a magazine  
quality  
a newspaper  
No way!

up-to-date  
prefer to  
thick  
a detail  
beat, beat, beaten  
old-fashioned  
out-of-date  
bother  
at any moment  
exactly  
go on the internet  
download

ihmeellinen, ihana  
keksintö  
viihdyttää  
seuralainen, kumppani  
tapahtuma  
piristää  
olla alla päin  
toisin kuin  
panna kiinni, sulkea  
panna päälle, avata  
milloin tahansa, aina kun  
huvittaa, haluta (tehdä jotakin)  
kanava  
valita  
dokumenttiohjelma  
eniten  
kauan eläköön...!

televisio-ohjelma  
roska(a)  
viettää, kuluttaa  
kauheus, kauhu  
väkivalta  
melkein  
aivot  
ainakin  
mielikuvitus  
kommunikoida,  
olla kanssakäymisissä  
luonnollisesti  
ihminen  
vierailija  
televio on lomaannuttanut  
heidät  
telkkari

valita jokin radioasema  
kuunnella jotakin  
varsinkin, erityisesti  
suosikki  
uutiset  
sää  
ajankohtaisohjelma  
lopulta, kuitenkin, joka tapauksessa  
täysi(n)  
typerys, idiootti  
lautapeli  
kaveri  
pikkusisko  
panna pois päältä, sammuttaa, sulkea  
harvoin  
kaupallinen; mainos  
enää  
törky, roska  
joko - tai  
älykääpiö  
nauraa jollekin  
typerä, tylsä  
vitsi  
ääliö, idiootti  
niin sanottu, muka  
tehda hulluksi  
ruokkia  
lukea  
arvostaa, ymmärtää  
luen kyllä  
aikakauslehti  
laatu  
sanomalehti  
Ei ikinä!

ajan tasalla, ajankohtainen  
pitää jostakin enemmän kuin  
paksu  
yksityiskohta  
voittaa  
vanhanaikainen  
vanhanaikainen  
viitsiä  
koska tahansa, minä hetkenä  
tahansa  
täsmälleen  
mennä nettiin  
imuroida, noutaa tiedostoja,

print out	ladata tietoverkosta
keep, kept, kept	tulostaa
a pice of news	säilyttää, pitää uutinen
no harm in it	haitta, vahinko
harm	mainonta
advertising	kuluttaja
a consumer	hyödyllinen
useful	tuote
a product	johtaa johonkin
lead, led, led to	luulla
suppose	mainos
an advert	informatiivinen, tietoa
informative	välittävä
of course	tietenkin, tietenkään
necessarily	välttämättä
believe	uskoa
common sense	terve järki
understandable	ymmärrettävä
a manufacturer	valmistaja
exaggerate	liioitella
a waste	hukka, tuhlaus
can't stand	ei voida sietää
a TV commercial	televisiomainos
space	tila
a website	www-sivusto, saitti
a pop-up	pop up -ikkuna, ponnahdusikkuna
a banner	banneri, mainospalkki
irritating	ärsyttävä
an ad	mainos
an advertiser	mainostaja
buy, bought, bought	ostaa
admit	myöntää
attractive	viehättävä, houkutteleva

### KEY ENGLISH 3

#### 27A Simple things you can do to help protect our environment

environment	ympäristö
energy	energia
a cooker	liesi
a heater	lämmityslaite
electric	sähkö-, sähköinen
burn	polttaa; palaa
oil	öljy
gas	kaasu
coal	kivihiili
turn off	sammuttaa
an appliance	laite
a fridge	jääkaappi
air	ilma
electricity	sähkö-, sähköinen
cool down	viilentää
fit	laittaa, asentaa
a lamp	lamppu
energy-efficient	energiaa säästävä
a light bulb	hehkulamppu
a washing machine	pyykinpesukone
a TV set	televisio (vastaanotin)
a model	malli
each of us	jokainen meistä
alitre	litra
a taker lorry	säiliöauto
run, ran, run	käyttää, pitää käynnissä
a dishwasher	astianpesukone
a load	lasti, pesukoneellinen
minimum	pienin mahdollinen
a flush	huuhtelu
dual-flush	kaksoishuutelu
a toilet	WC, vessa
waste	jätteet, jäte
a drink can	juomatölkki
be worth of	olla jonkin arvoinen
a jar	tölkki, tuoppi
food scraps	ruoantähteet
plastic	muovi
a container	säiliö, astia
soil	maaperä

forever  
recycle  
a forest  
a recycling bank  
organic  
vegetable peelings  
a basket  
save  
a bin

aina, ikuisesti  
kierrättää, käyttää uudelleen  
metsä  
kierrätyspiste  
orgaaninen  
vihannesten kuoret  
kori  
säästää, pelastaa  
roska-astia

### KEY ENGLISH 3 27B The state of the planet

a state  
a planet  
the planet  
the Earth  
cause  
damage  
breathe  
be polluted  
a rainforest  
chop down  
die out  
use up  
resources  
fossil fuels  
main points

a species, (mon.) species  
the equator  
a scientist  
estimate  
up to  
die out  
officially  
classify as

endangered

pollution  
produce  
an amount  
break down  
harmless  
soil  
contain  
a poison  
pollute

depend on  
a factory  
a vehicle  
provide  
fossil fuels  
acid rain  
add to  
global warming  
nuclear power  
produce  
a hazard  
radiation  
a leak  
nuclear waste  
natural resources  
renewable  
energy conservation  
alternative  
a solar panel  
a wind farm  
a windmill

the Ice Age  
since  
in the future  
a temperature  
rise, rose, risen  
cause  
a blanket  
the greenhouse effect  
trap  
heat

tila, tilanne  
planeetta, kiertotähti  
maapallo, maailma  
maa(pallo), maailma  
aiheuttaa  
vahinko, vaurio  
hengittää  
saastua  
sademetsä  
kaataa (puu)  
kuolla sukupuuttoon  
kuluttaa loppuun  
luonnonvarat  
fossiiliset polttoaineet  
pääkohdat

laji  
päiväntasaaja  
tiedemies  
arvioida  
(tässä:) jopa  
kuolla sukupuuttoon  
virallisesti  
luokitella johonkin  
vaarassa kuolla  
sukupuuttoon,  
uhanalainen

saasteet, saastuminen  
tuottaa, saada aikaan  
määrä  
hajota  
vaaraton, harmiton  
maaperä  
sisältää  
myrkky  
saastuttaa

luottaa johonkin  
tehdas  
kulkuväline  
tuottaa, tarjota  
fossiiliset polttoaineet  
(esim. hiili ja öljy)  
hapan sade, happosade  
lisätä johonkin  
maapallon lämpeneminen  
ydinvoima  
tuottaa  
vaara, riski  
säteily  
vuoto  
ydinjäte  
luonnonvarat  
uusiutuva  
energian säästäminen  
vaihtoehtoinen  
aurinkopaneeli  
tuulivoimala  
tuulimylly

jääkausi  
jostakin lähtien  
tulevaisuudessa  
lämpötila  
nousta  
aiheuttaa  
peite, kerros  
kasvihuoneilmiö  
vangita  
lämpö

thus  
carbon dioxide  
a cause  
a scientist  
upset  
a flood  
a shortage  
a storm  
warm up  
a polar ice cap  
melt  
a level

näin, täten  
hiilidioksidi  
syy  
tiedemies  
häiritä, järkyttää  
tulva  
puute, pula  
myrsky  
lämmetä  
napajää(tikkö)  
sulaa  
pinta

## SPOTLIGHT 1 1A On the net

the net  
live  
Southern Finland  
Finnsih  
English  
often  
miss  
a friend  
a primary school  
keep together  
keep, kept  
different  
a country  
now

still  
keep in touch  
on holiday  
You bet.  
lovely  
hear, heard  
like  
in August  
a few  
left  
How are things in Scotland  
quiet  
a village  
called  
exciting  
an autumn  
come, came  
visit  
weird  
be, was/were  
as you know  
know, knew  
an eend  
a winter  
pretty  
hot  
go swimming  
go, went  
have, had  
a beach party  
tonight  
Lucky you.  
cold  
cool  
a hospital  
serious  
an operation  
on Tuesday  
a doctor  
say, said  
will be  
a beginning  
sure  
make friends  
make, made  
lots of  
called  
life  
the States  
Howdy  
everybody  
friendly  
a smile

netti  
asua  
Etelä-Suomi  
suomalainen  
englantilainen  
usein  
kaivata  
ystävä  
alakoulu  
pitää yhdessä  
pitää yhdessä  
eri, erilainen  
maa  
nyt  
vielä,  
yhä  
pitää, yhteyttä  
lomalla  
Taatusti  
ihana  
kuulla  
kuten, niin kuin  
elokuussa  
muutama  
jäljellä  
Miten Skotlannissa menee ?  
hiljainen  
kylöä  
nimeltä  
jännittävä  
syksy  
tulla  
vieraila  
outo, kummallinen  
olla  
kuten tiedät  
tietää, tuntea  
loppu  
talvi  
melko  
kuuma  
käydä uimassa  
mennä  
olla jollakin  
rantajuhlat  
tänä iltana  
Onnenpekka!  
kylmä  
kiva, mahtava  
sairaala  
vakava  
leikkaus  
tiistaina  
lääkäri  
sanoa  
on (tulevaisuudessa)  
alku  
varma  
saada ystäviä  
tehdä, valmistaa  
paljon  
kutsua, soittaa  
elämä  
Yhdysvallat  
hei, moi  
jokainen, kaikki  
ystävällinen  
hymy

talk  
 plenty of  
 food  
 too  
 a pupil  
 over  
 a thousand  
 even  
 everything  
 feel, felt  
 a little  
 lonely  
 sometimes  
 luckily  
 Yoy got it.  
 worry  
 Things will work out.  
 true  
 a wildlife picnic  
 tomorrow  
 wild  
 Don't get too close

puhe; puhua  
 paljon, runsaasti  
 ruoka  
 liian  
 oppilas  
 yli  
 tuhat  
 jopa  
 kaikkia  
 tuntea (olonsa)  
 vähän  
 yksinäinen  
 joskus  
 onneksi  
 Sinä sen sanoit.  
 huolehtia, olla huolissaan  
 Asiat järjestyvät.  
 totta  
 luontoretki  
 huomenna  
 villi, raju  
 älä mene liian lähelle

### SPOTLIGHT 1 3 @School

#### 3A Question time

elementary school (AmE)  
 a grade (AmE)  
 high school  
 primary school (BrE)  
 preschool  
 basic education  
 lower secondary  
 a system

alakoulu  
 luokka-aste  
 yläkoulu ja lukio  
 alakoulu  
 yläkoulu ja lukio  
 esikoulu  
 yläkoulu  
 järjestelmä

the united States of America  
 a town  
 a Finn  
 has been  
 a month  
 because  
 a student  
 has given a talk  
 a classmate  
 has explained  
 a school subject  
 a rule  
 a schedule  
 a question

Yhdysvallat  
 kaupunki  
 suomalainen (henkilö)  
 on ollut  
 kuukausi  
 koska  
 oppilas, opiskelija  
 on pitänyt puheen  
 luokkakaveri  
 on selittänyt  
 kouluaine  
 sääntö  
 lukujärjestys  
 kysymys

#### 3B Homework on homework

have to  
 \*make, made, made  
 a poster  
 around the world  
 email  
 greetings  
 busy  
 science  
 finish  
 an essay  
 literature  
 the movies  
 a project  
 yet  
 altogether  
 \*spend, spent, spent  
 an hour  
 remember  
 used to  
 complain  
 love=(with) love  
 a way  
 we've had  
 discussion  
 a class  
 \*do homework  
 have they been  
 demanding  
 a test

täytyä  
 tehdä, valmistaa  
 poster  
 ympäri maailmaa  
 lähettää sähköpostia  
 terveiset  
 kiireinen  
 luonnontieteet  
 saada loppuun  
 kirjoitelma  
 kirjallisuus  
 elokuvat  
 projekti  
 vielä  
 kaiken kaikkiaan, yhteensä  
 viettää, kuluttaa aikaa  
 tunti  
 muistaa  
 oli tapana  
 valittaa  
 terveisin  
 tapa  
 meillä on ollut  
 keskustelu  
 luokka  
 tehdä kotitehtävät  
 ovatko he olleet  
 vaativa  
 koe

mostly  
multiple-choice  
I haven't written  
excellent  
how come  
I've used  
of course  
we have spoken  
have you studied  
study  
a language  
French  
German  
learn  
important  
I've started  
Spanish  
a lesson  
have you read  
a hundred  
I haven't counted  
lately  
I haven't had  
busy  
after-school activities  
What have you taken up?  
I've become  
I've joined  
basketball  
a team  
friendly  
supportive  
miss  
relaxed  
for example  
call  
a first name  
written permission  
a break  
sound  
laid-back  
you've told  
interesting  
a vacation  
since

enimmäkseen  
monivalinta-  
en ole kirjoittanut  
erinomainen  
miksi  
olen käyttänyt  
tietysti  
olemme puhuneet  
oletko opiskellut  
opiskella  
kieli  
ranskan kieli  
saksan kieli  
oppia, opiskella  
tärkeä  
olen aloittanut  
espanjan kieli  
oppitunti  
oletko lukenut  
sata  
en ole laskenut  
viime aikoina  
minulla ei ole ollut  
kiireinen  
koulun kerhotoiminta  
Mitä olet aloittanut?  
minusta on tullut  
olen liittynyt  
koripallo  
joukkue  
ystävällinen  
kannustava  
kaivata  
rento  
esimerkiksi  
kutsua  
etunimi  
kirjallinen lupa  
välitunti, tauko  
kuulostaa  
rento  
olet kertonut  
mielenkiintoinen  
loma  
koska

## SPOTLIGHT 2

### 4A The L.A. Street Lions

\*get around  
a vehicle  
a passport  
a flight  
a ticket  
a passenger  
a station  
baggage  
travel  
nearby  
on foot  
belong to  
a dance group  
\*take part in  
an event  
coast to coast  
across  
altogether  
a student  
reserve  
a journey  
change  
decide  
a stopover  
the Windy City  
perform  
burn  
fuel  
harm  
an environment  
Greyhound  
\*get sick  
dizzy  
dumb  
agree with

liikkua paikasta toiseen  
kulkuneuvo  
passi  
lento  
lippu  
matkustaja  
asema  
matkatavarat  
matkustaa  
lähellä oleva  
jalkaisin  
kuulua  
tanssiryhmä  
ottaa osaa  
tapahtuma  
rannikolta toiselle  
halki  
yhteensä  
oppilas, opiskelija  
varata  
matkatavarat  
vaihtaa  
päättää  
pysähdys  
Chicagon lempinimi  
esiintyä  
polttaa  
polttoaine  
vahingoittaa  
ympäristö  
bussiyhtiön nimi  
voida pahoin  
pyörryksissä  
tyhmä, typerä  
olla samaa mieltä

a dining car  
 \*spend, spent, spent  
 a conductor  
 what if  
 \*get off  
 That's all there is to it.  
 a pillow  
 luckily  
 share  
 a seat  
 a toilet  
 a shower  
 a lack of room  
 own  
 either  
 relax  
 pull  
 \*become, became, become  
 No sweat!  
 upper  
 lower  
 mind  
 Deal?  
 Fine by me.  
 I wonder how  
 as is often the case  
 far  
 \*get hurt  
 a surprise  
 \*let go of  
 sprain  
 an ankle  
 a fault  
 maybe  
 a bruise  
 \*take off  
 a sock  
 \*have a look  
 swollen  
 an ice cube  
 a packet  
 a first-aid kit  
 a restroom  
 a bandage  
 a tape  
 almost  
 \*hold on to  
 serious  
 broken  
 a bone  
 blame  
 gonna = going to  
 miserable  
 a beginning  
 a trip of a lifetime

ravintolavaunu  
 kuluttaa, viettää  
 konduktööri  
 entä jos  
 nousta kyydistä  
 Se on sillä selvä.  
 tyyny  
 onneksi  
 jakaa  
 istuin  
 vessa  
 suihku  
 tilan puute  
 oma  
 kumpikaan  
 rauhoittua  
 vetää  
 tulla joksikin  
 Helppoa!  
 ylempi  
 alempi  
 panna pahakseen  
 Sopiiko?  
 Sopii minulle.  
 Mitenkähän  
 niin usein käy  
 pitkälle  
 satuttaa itsensä  
 yllätys  
 päästää irti  
 nyrjäyttää  
 nilkka  
 syy, vika  
 ehkä  
 mustelma  
 riisua  
 sukka  
 katsoa, vilkaista  
 turvonnut  
 jääpala  
 tasku  
 ensiapulaukku  
 vessa  
 side(harso)  
 teippi  
 melkein  
 pitää kiinni  
 vakava  
 murtunut, rikki  
 luu  
 syyttää  
 aikoa  
 surkea  
 alku  
 elämänsä matka

## **SPOTLIGHT 2** **2A Action weekend**

paint  
 a model  
 quiet  
 \*be into action sports  
 stuff  
 a few  
 \*hang out  
 worried about  
 an art course  
 for years  
 learn  
 quite  
 a coach  
 nowadays  
 decide  
 cheap  
 expensive  
 recommend  
 a campfire  
 exciting  
 an outdoor activity centre  
 a group leader  
 a youngster

maalata  
 pienoismalli  
 hiljainen  
 harrastaa toimintalajeja  
 jutut, asiat  
 muutama  
 viettää aikaa, olla  
 huolissaan  
 taidekurssi  
 vuosikausi  
 oppia  
 melko  
 valmentaja  
 nykyään  
 päättää  
 halpa  
 kallis  
 suositella  
 leirinuotio  
 jännittävä  
 leirikeskus  
 ryhmänohjaaja  
 nuori

attend  
 popular  
 \*find, found, found  
 hard  
 \*take part in  
 loads of  
 an activity  
 early  
 a meal  
 worry about  
 lonely  
 belong to  
 caving  
 a helmet  
 hardly  
 breathe  
 dark  
 support  
 encourage  
 \*leave, left, left  
 a mobile  
 glasses  
 a cave  
 a spare pair  
 proud of  
 complain about  
 pretend  
 brave  
 \*be scared to death  
 \*let, let, let  
 \*show, showed, shown  
 face  
 a fear  
 \*give up  
 \*give in  
 plenty  
 in the first place  
 quiet  
 \*draw, drew, drawn  
 wish  
 at least  
 trekking  
 abseiling  
 rock climbing  
 You must be joking.  
 a difference  
 \*split, split, split  
 a hill  
 in a good mood  
 \*lead, led, led  
 an aim  
 a castle  
 \*keep up  
 stay  
 cheerful  
 tired  
 a record  
 photography  
 a packed lunch  
 we made our way  
 silly  
 I began to enjoy myself  
 at last  
 equipment  
 several  
 tough  
 a rope  
 a couple of  
 a harness  
 a huge amount of  
 courage  
 an instructor  
 \*be in control  
 dangerous  
 a cliff  
 a foot, feet  
 safety  
 believe

ottaa osaa  
 suosittu  
 olla jotakin mieltä, löytää  
 raskas  
 ottaa osaa  
 monta, paljon  
 aktiviteetti, toiminto  
 aikainen  
 ateria  
 huolehtia  
 yksinäinen  
 kuulua  
 luolakiipeily  
 kypärä  
 tuskin  
 hengittää  
 pimeä  
 tukea  
 rohkaista  
 jättää, lähteä  
 kännykkä, silmälasit  
 silmälasit  
 luola  
 varapari  
 ylpeä  
 valittaa  
 teeskennellä  
 rohkea  
 pelätä kuollakseen  
 antaa, sallia  
 näkyä, näyttää  
 kohdata  
 pelko  
 antaa periksi  
 luovuttaa  
 runsaasti  
 alun alkuaan  
 hiljainen  
 piirtää  
 toivoa  
 ainakiin  
 vaeltaminen  
 köysilaskeutuminen  
 kalliokiipeily  
 Et voi olla tosissasi.  
 ero  
 jakaantua, jakaa  
 mäki, kukkula  
 hyvällä tuulella  
 johtaa  
 tavoite  
 linna  
 pysyä mukana  
 pysyä  
 hyväntuulinen  
 väsynyt  
 ennätys  
 valokuvaus  
 eväät  
 pääsimme  
 typerä, tyhmä  
 aloin viihtyä  
 lopulta  
 varusteet  
 useat  
 kestävä  
 köysi  
 pari  
 valjaat  
 valtava määrä  
 rohkeus  
 ohjaaja  
 olla vastuussa, määrätä  
 vaarallinen  
 kalliokiipeily  
 jalka, jalat  
 turvallisuus  
 uskoa

### **SPOTLIGHT 3** **3A Media mania**

the media

tiedotusvälineet, media



the press  
 a newspaper  
 a freesheet  
 a tabloid  
 a quality paper  
 a magazine  
 a channel  
 a show  
 a series  
 a drama  
 a documentary  
 a quiz show  
 a reality show  
 the weather forecast  
 news  
 a web page  
 a website  
 social media  
 a discussion forum  
 share  
 a viewer  
 a journalist  
 a newsreader  
 a sports commentator  
 a host  
 a presenter  
 a gadget  
 a PC (personal computer)  
 a smartphone  
 use  
 at work  
 a sports event  
 even  
 on public transport  
 experience  
 endless  
 a stream  
 information  
 entertainment  
 \*deal with  
 deal, dealt, dealt  
 a way  
 watch  
 on my computer  
 in front of  
 a telly  
 certain  
 hooked on  
 \*spend, spent, spent  
 at least  
 a box  
 an advert  
 \*be on  
 listen to  
 on the radio  
 used to  
 a programme  
 bother  
 \*choose, chose, chosen  
 instead of  
 on the whole  
 apart from  
 \*mean, meant, meant  
 which  
 on the go  
 on my phone  
 \*buy online  
 illegal  
 necessary  
 printed  
 regularly  
 \*feel, felt, felt  
 on TV  
 on the internet  
 pick up  
 for free  
 \*have access to  
 either-or  
 mostly  
 for fun  
 as  
 on social media  
 interact with  
 text  
 a source

lehdistö  
 sanomalehti  
 ilmaislehti  
 iltapäivälehti  
 laatu-lehti  
 aikakauslehti  
 kanava  
 ohjelma  
 (televisio)sarja  
 draama, näytelmä  
 dokumentti  
 visailuohjelma  
 tosi-tv-ohjelma  
 säätiedotus  
 uutiset  
 nettisivu  
 nettisivusto  
 sosiaalinen media  
 keskustelupalsta  
 jakaa  
 (television)katselija  
 toimittaja  
 uutistenlukija  
 urheiluselostaja  
 juontaja  
 juontaja  
 laite  
 tietokone  
 älypuhelin  
 käyttää  
 työssä  
 urheilutapahtuma  
 jopa  
 julkisissa liikennevälineissä  
 kokea  
 loputon  
 virta  
 tieto, informaatio  
 viihde  
 käsitellä  
 jakaa  
 tapa  
 katsella  
 tietokoneeltani  
 edessä  
 telkkari, tv  
 tietty  
 koukussa  
 viettää, kuluttaa  
 ainakin  
 telkkari  
 mainos  
 olla menossa  
 kuunnella  
 radiosta, -ssa  
 oli tapana  
 ohjelma  
 vaivautua, viitsiä  
 valita  
 sen sijaan, että  
 kaiken kaikkiaan  
 paitsi, lukuun ottamatta  
 tarkoittaa  
 joka, mikä  
 liikkeellä  
 puhelimestani, puhelimellani  
 ostaa netistä  
 laiton  
 tarpeellinen  
 painettu  
 säännöllisesti  
 tuntea, olla mieltä  
 televisiosta, -ssa  
 internetistä, -ssä  
 poimia, ottaa  
 ilmaiseksi  
 olla yhteys  
 joko-tai  
 useimmiten, pääasiassa  
 huvia, huviksi  
 koska  
 sosiaalisessa mediassa  
 olla tekemisissä  
 lähettää tekstiviestejä  
 lähde

a topic  
 particularly  
 a project  
 a fact  
 fiction  
 show  
 amazing  
 a video clip  
 a penguin  
 the Antarctic  
 a presenter  
 \*take to the air  
 across  
 icy  
 an ocean  
 land  
 a rainforest  
 \*go on  
 explain  
 dislike  
 April Fool's Day  
 \*come up with  
 \*come, come, come  
 an all-time classic  
 Swiss  
 harvest  
 \*grow, grew, grown  
 politely  
 place  
 a tin  
 a sauce  
 \*get angry  
 fool  
 a producer  
 should  
 realize  
 trust  
 a content  
 acces  
 reliable  
 informative  
 misleading  
 a lie  
 double-check  
 common sense

aihe  
 erityisesti  
 projekti  
 tosiasia, fakta  
 taru, kaunokirjallisuus  
 näyttää, esittää  
 hämmästyttävä  
 video  
 pingviini  
 Etelämanner, Antarktis  
 juontaja  
 lähteä lentoon  
 poikki, halki  
 jäinen  
 valtameri  
 laskeutua  
 sademetsä  
 jatkaa  
 selittää  
 ei pitää  
 aprillipäivä  
 keksiä, löytää  
 tulla  
 kaikkien aikojen klassikko  
 sveitsiläinen  
 sadonkorjuu; korjata satoa  
 kasvattaa, kasvaa  
 kohteliaasti  
 panna, asettaa  
 säilyketölkki  
 kastike  
 suuttua  
 puijata, huijata  
 tuottaja  
 pitäisi  
 tajuta  
 luottaa  
 sisältö  
 päästä käsiksi  
 luotettava  
 asiapitoinen, informatiivinen  
 harhaanjohtava  
 valhe  
 tarkistaa uudelleen  
 terve järki

### SPOTLIGHT 3

#### 5A Open your eyes

pollution  
 a solution  
 a nature lover  
 protect  
 the environment  
 reduce  
 reuse  
 recycle  
 a choice  
 save  
 waste  
 the earth  
 unwanted  
 the ground  
 a substance  
 harm  
 cover  
 nature  
 disturb  
 wildlife  
 offer  
 \*put up  
 a tent  
 a landowner  
 permission  
 sunblock  
 a footprint  
 \*may  
 next to  
 appreciate  
 \*take for granted  
 \*take care of  
 \*cut down on  
 sort

saaste  
 ratkaisu  
 luonnonystävä  
 suojella  
 ympäristö  
 vähentää  
 käyttää uudelleen  
 kierrättää  
 valinta  
 pelastaa  
 tuhjata; jäte  
 maapallo  
 ei-toivottu  
 maapallo  
 aine  
 vahingoittaa  
 peittää, kaataa  
 luonto  
 häiritä  
 luonto, villieläimet ja  
 kasvit  
 tarjota  
 pystyttää  
 teltta  
 maanomistaja  
 lupa  
 aurinkovoide  
 jalanjälki  
 saattaa  
 vieressä, lähellä  
 arvostaa  
 pitää itsestään selvänä  
 huolehtia  
 vähentää  
 lajitella

rubbish  
\*make  
wonderful  
an experience  
look after  
\*be able to  
\*be proud of  
\*be aware of  
including  
cycle  
glass  
cardboard  
a can  
bio waste  
electricity  
switch off  
the least  
actually  
experience  
breathtaking  
could have spent  
the whole  
\*go wild camping  
a campfire  
at night  
visit  
Mount  
Sicily  
a volcano  
a heart  
thick  
sulphur  
bubble  
smell  
if I had stayed  
would have fainted  
amazing  
below  
the future  
a rainforest  
a biome  
Mediterranean  
a plant  
extraordinary  
unique  
even  
covered in class  
could have fitted  
huge  
should have gone  
\*be amazed by  
a fir tree  
a birch tree  
a pine tree  
a blueberry  
purity  
a lack of  
healthy  
complain about  
report on  
the Lake District  
camping  
organized  
a campsite  
the wilds  
\*bring  
a responseability  
toilet waste  
ugly  
a threat  
a hole  
(a)grass  
\*make fire  
as  
a trekker  
a wild camper  
look forward to  
at times  
\*be disgusted by  
behavior  
obviously  
care about  
I couldn't help but  
compare  
seem

roskat, jätteet  
saada(tekemään)  
ihmeellinen, ihana  
kokemus  
pitää huolta, huolehtia  
pystyä  
olla ylpeä  
olla tietoinen  
mukaan lukien  
pyörällä  
lasi  
pöytä, kartonki  
tölkki  
biojäte  
sähkö  
sammuttaa  
vähiten, vähintään  
itse asiassa, tosiaankin  
kokea  
henkeäsalpaava  
olisi voinut viettää  
koko  
mennä retkeilemään  
nuotio  
yöllä, öisin  
vierailu  
vuori(erinimisissä)  
Sisilia  
tulivuori  
sydän, ydin  
paksu  
rikki  
pulput, kuplia  
haista, haistaa  
jos olisin jäänyt  
olisi pyörtynyt  
hämmentävä  
alapuolella, ala  
tulevaisuus  
sademetsä  
eloyhteisö, biomi  
välimerellinen  
kasvi  
erikoinen  
ainutlaatuinen  
jopa  
lasikaton alla  
olisi voinut mahtua  
valtava  
olisi pitänyt mennä  
olla hämmästynt  
kuusi  
koivu  
mänty  
mustikka  
puhtaus  
puute, pöytä  
terveellinen  
valittaa  
raportoida  
Englannin järvialue  
telttailu  
järjestetty  
leirintäalue  
erämaa, syrjäinen alue  
tuoda  
velvollisuus, vastuu  
käymäläjäte  
ruma  
uhka  
reikä  
ruoho, nurmikko  
sytyttää nuotio  
koska  
patikoija  
retkeilijä  
odottaa innokkaana  
ajoittain  
inhottaa  
käytös  
ilmeisesti  
välittää, huolehtia  
en voinut olla vertaamatta  
näyttää

natural

few

in fact

a human being

\*make camp

behave

however

a spot

vegetation

a site

a memory

luonnonläheinen,

luonnollinen

harvoja

itse asiassa

ihminen

leiriytyä

käyttäytyä

kuitenkin

paikka, kohta

kasvillisuus

paikka

muisto

## Appendix II

Queries and tags used in the corpus study.

SYNTAX	MEANING
word	exact word
[lemma]	all forms of a word
word.[pos]	exact word and part of speech
-word /[pos]	not followed by a word or PoS
*	wildcard i.e. any word or form
	or
N	noun
NN	common noun
NP	proper noun
J	adjective
V	verb
UH	interjection
PPY	2nd person personal pronoun (you)
RR	general adverb

QUERIES		
kid.[nn*]	laid back	(the box)
city slicker.[n*]	granny.[n*]	search: [=tv]
dad.[n*]	kid.[nn*]	airhead.[nn*]
Mind you!	gran.[n*]	pal.[nn*]
search: mind.[vv*] [PPY] -[V*]	a stomach bug	kid sister.[nn*]
a bit	[crack] a joke	a jerk
search: a.[RR21] bit.[RR22]	knackered	No way!
guy.[nn*]	gonna = going to	search: no way !
hick.[nn*]	search: gon na	ad.[nn*]
buzz off	(a Yankee)	wacko.[nn*]
(pest.[nn*] )	a bit of	ad=advert=advertisement
goalie.[nn*]	guy.[nn*]	search: ad.[nn*]
You bet	the States.[np*]	advert.[nn*]
the States.[np*]	(Down Under)	advertisement.[nn*]
howdy.[UH]	(mad)	acquire.[vv*]
crush.[nn*]	crazy.[j*]	porn.[nn*]
pretty -[nn*]	I guess	thus
luv=love	footy	telly.[nn*]
search: .! : luv , ;	(globetrotter)	(a box)
cussing.[nn*]	Aussie	on the go
mate	stuff	vegetation.[nn*]
how come		

## Appendix III

All the seventy-nine stylistically restricted words and their translations found in the eighteen chapters examined for the study are listed in the tables below. The abbreviations used in the columns for the three dictionaries (OALD, COBUILD and MEDAL) refer to the word class specification the headword in question carries, and to the labels assigned to the words. The numbers after the word class specifications refer to the numbers used in the dictionaries to distinguish the homonyms/ word senses from one another. Number zero (0) is used if the word or phrase does not occur in the dictionary in the sense intended and a hyphen (-) is used if no labels or other relevant extra information occur in connection with the word in the dictionary. Other abbreviations used in the tables are explained below.

Abbreviation	Meaning
IDM	idiom
BrE	British English
AmE	American English
NAmE	North American English
AusE	Australian English
NZE	New Zealand English

Abbreviation	Word class
N	noun
V	verb
ADJ	adjective
ADV	adverb

Abbreviation	Label
F	Formal
IF	Informal
S	Slang

Grade 7, Theme I: Immediate surroundings						
Book	Chapter (Text)	Word	Translation	OALD	COBUILD	MEDAL
<i>Key English I</i>	1 (1A)	a kid	lapsi, nuori	N 1 IF	N 1 IF	N 1 IF
		a city slicker	kaupunkilainen, juppi	N IF	N IF	N IF
		dad	isä	N IF	N IF	N IF
		Mind you!	Ota huomioon!, Muuten...	V IDM IF	2 MIND PHRASES -	(2)V PHRASES -
		a bit	hieman	ADV- esp. BrE	ADV -	(1) ADV, Pron PHRASES 1 IF
		a guy	tyyppi, jätkä	N 1 IF	N 1 IF	N 1 IF
		a hick	juntti	N IF, esp. NAmE	N IF	N mainly AmE IF
		buzz off	painua tiehensä	BUZZ V PHRASAL V IF	0	(1) V PHRASAL V IF
<i>Smart Moves I</i>	2	pest	kiusankappale	N 2 IF	N 2 IF	N 2 IF
		cute	söpö, suloinen	ADJ 1 -	ADJ 1 IF	ADJ 1 -
		goalie	maalivahti	GOALKEEPER N IF	N IF	N IF
<i>Spotlight I</i>	1 (1A)	the net	netti	N 5 IF	N 4 -	(1) N 3 -
		You bet	Taatusti	BET V IDM 2 IF	BET V PHRASES 13 IF	BET (1) V PHRASES -
		Lucky you!	onnenpekka!	LUCKY ADJ IDM IF	0	LUCKY ADJ PHRASES -
		cool	kiva, mahtava	ADJ 5 IF	ADJ 7 -	(1) ADJ 4 -
		the States	Yhdysvallat	N 6 IF	N 3 IF	STATE (1) N 4 IF
		Howdy	hei, moi	NAmE IF	IF	IF

Grade 7, Theme II: School						
Book	Chapter	Word	Translation	OALD	COBUILD	MEDAL
<i>Key English I</i>	3	cool	viileä, upea, hieno	ADJ APPROVING 5 IF	ADJ 7 -	(1) ADJ 4 -
		crush	ihastus	0	6 N IF	(2) N 2 IF
		cute	söpö, suloinen	ADJ 1 -	ADJ 1 IF	ADJ 1 -
		pretty	aika, melko	ADV rather IF	ADV 3 IF	(1) ADV 1 -
		luv=love	terveisin	N IF	IF BrE	N BrE IF
		cussing	kiroilu	CUSS V IF	CUSS IF	CUSS V mainly AmE IF
<i>Smart Moves I</i>	12	mate	kaveri	N1,2 BrE, AustralE, IF	N 1 BrE, IF	(1) N 1 BrE IF
<i>Spotlight I</i>	3 (3A, 3B)	how come	miksi	HOW IDM 2 IF	HOW PHRASES 13 IF	HOW ADV, Conj PHRASES -
		laid back	rento	ADJ IF	ADJ IF	ADJ IF
		luv=(with)love	terveisin	N IF	IF BrE	N BrE IF
		tricky	hankala, vaikea	ADJ 1 rather IF	ADJ -	ADJ 1 -

Grade: 8, Theme I: Travelling						
Book	Chapter	Word	Translation	OALD	COBUILD	MEDAL
<i>Key English I</i>	10	a granny	mummu	N IF	N IF	N IF
		a kid	lapsi, nuori	N 1 IF	N 1 IF	N 1 IF
		gran	mummu	BrE, IF	N BrE, IF	N IF
		a (stomach) bug	vatsatauti	N 2 IF	BUG N 2 IF	BUG (1) N 1 IF
<i>Smart Moves II</i>	17	crack a joke	kertoa, 'murjaista' vitsi	V 10 IF	CRACK V 8 -	CRACK (1) V PHRASES IF
		knackered	poikki, 'rättiväsynyt'	ADJ BrE S	ADJ 1 BrE, IF	ADJ 1 BrE S
		sleep like a log	nukkua kuin tukki	V IDM IF	0	SLEEP (1) V PHRASES -
<i>Spotlight II</i>	4 (4A)	No sweat!	helppoa!	N IDM IF	0	SWEAT (1) N PHRASES -
		gonna=going to	aikoa p. 89 PUHEESSA	IF	IF	IF

Grade: 8, Theme II: Free Time & Hobbies						
Book	Chapter	Word	Translation	OALD	COBUILD	MEDAL
Key English II	13 (13 B)	a whole new ball game	(myös: aivan uusi asia)	N BALL GAME IDM IF	BALL GAME PHRASE -	BALL GAME PHRASE -
		a Yankee	jenkki	N BrE, IF	N 2 IF	N 1 -
		dozens of	kymmeniä, suuri määrä	DOZEN N 3 (pl.) IF	DOZEN N 3 -	DOZEN 2 -
		a bit of	melkoinen, aikamoinen	BIT N 5 IF, esp. BrE	0	BIT PHRASES IF
		a guy	mies, kaveri, tyyppi	N 1 IF	N 1 IF	N 1 IF
		the States	Yhdysvallat	STATE N 6 IF	STATE N 3 IF	STATE N 4 IF
Smart Moves II	1	Down Under	Australia (ja Uusi-Seelanti)	DOWN IDM IF	mainly BrE, IF	N IF
		cute	söpö	ADJ 1 -	ADJ 1 IF	ADJ 1 -
		mad	hullu	ADJ 2 IF, esp. BrE	ADJ -	ADJ 1a IF, 1b IF
		crazy	hullu	ADJ 1 IF, esp. NAmE	ADJ 1 IF	ADJ IF
		I guess	(tässä:) varmaankin	GUESS V 3 IF, esp. NAmE	GUESS PHRASES 6 mainly AmE, IF	GUESS PHRASES IF
		footy	australialainen jalkapallo	N BrE, IF (=football)	FOOTIE (footy) N BrE IF (=football)	0 FOOTIE BrE IF (= the game of football)
		globetrotter	maailman-matkaaja	GLOBETROTTING ADJ IF N -	GLOBETROT IF (-trotter -)	N IF
		G'day	hei, tervehdys (austr.)	AusE, NZE	0	AusE IF
		Take care!	pärjää! / jaksaa!	CARE N IDM IF	CARE PHRASES 15 -	CARE PHRASES TAKE CARE 2 -
		Aussie	australialainen	N IF	ADJ IF	N IF
Spotlight II	2 (2A)	stuff	jutut, asiat	N IF	N 1 IF	N IF
		youngster	nuori	N IF	N -	N -
		you must be joking	et voi olla tosissasi	JOKE IDM IF	JOKE PHRASES 11 -	JOKE PHRASES -



Grade: 9, Theme I: Media						
Book	Chapter	Word	Translation	OALD	COBUILD	MEDAL
Key English III	26 (26A)	rubbish	roska(a)	N-, ADJ BrE IF	N 1 mainly BrE	N 1 -
		the box	telkkari	N 6 IF, esp. BrE	N 5 BrE, IF	(1) N BrE IF
		an airhead	typerys	N IF	N IF	N IF
		a pal	kaveri	N 1 IF, becoming old-fashioned	N IF, old-fashioned	N 1 IF old fashioned
		a kidsister	pikkusisko	KID A IF esp. NAmE	KID 2 IF	KID (3) A IF
		corny	typerä, tylsä	ADJ IF	ADJ -	ADJ -
		a jerk	ääliö, idiootti	N 2 IF	2 N IF	(2) N 2 -
		No way!	Ei ikinä!	N IDM IF	WAY PHRASES 57 IF	WAY (1) N 1 PHRASES -
		an ad	mainos	N IF	N IF	N IF
Smart Moves III	23	wacko	hullu	N IF	ADJ IF	ADJ IF N -
		spamming	roskapostin lähettäminen	N IF	SPAM V COMPUTING	N COMPUTING
		spam	roskaposti	N 2 IF	SPAM V COMPUTING	N COMPUTING
		ad=advert=advertisement	mainos	N IF / N BrE / N 1 -	N IF / N BrE / N written	N IF / N BrE / N-
		acquire	saada haltuunsa	V F	V 1 F	V 1 F
		porn	porno	N IF	N IF	N IF
		thus	täten	ADV F	ADV F	ADV F
Spotlight III	3 (3A)	a telly	telkkari, tv	N 1 BrE, IF	N BrE, IF	N BrE IF
		a box	telkkari	N 6 IF, esp. BrE	N 5 BrE, IF	(1) N BrE IF
		on the go	liikkeellä	N IDM IF	3 GO PHRASES 19 IF	GO (2) N PHRASES IF
		dislike	ei pitää	V rather F	V -	(1) V -

Grade: 9, Theme II: Sustainable Development						
Book	Chapter	Word	Translation	OALD	COBUILD	MEDAL
Key English III	27(27A,B)	an appliance	laite	N -	N 1 F	N -
		a vehicle	kulkuväline	N rather F	N 1 -	N -
Spotlight III	5 (5A)	vegetation	kasvillisuus	N -	N F	N F

## Appendix IV

Results of the corpus study. Raw and standardised frequencies of the words searched in COCA. An asterisk (\*) is placed before the words whose frequencies were counted manually.

<i>Word</i>		ALL	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC
<i>kid</i>	FREQ	36,058.00	8,820.00	13,727.00	6,641.00	6,308.00	562.00
	PER MIL	77.65	92.29	151.80	69.50	68.78	6.17
<i>city slicker</i>	FREQ	37.00	4.00	12.00	12.00	9.00	0.00
	PER MIL	0.08	0.04	0.13	0.13	0.10	0.00
<i>dad</i>	FREQ	40,873.00	8,449.00	18,535.00	7,594.00	5,672.00	623.00
	PER MIL	88.02	88.41	204.97	79.47	61.84	6.84
<i>Mind you!</i>	FREQ	742.00	146.00	278.00	178.00	114.00	26.00
	PER MIL	1.60	1.53	3.07	1.86	1.24	0.29
<i>a bit</i>	FREQ	26,354.00	5,502.00	8,533.00	6,308.00	4,606.00	1,405.00
	PER MIL	56.76	57.57	94.36	66.01	50.22	15.43
<i>guy</i>	FREQ	77,105.00	28,734.00	20,751.00	13,910.00	12,958.00	752.00
	PER MIL	166.05	300.67	229.47	145.56	141.28	8.26
<i>hick</i>	FREQ	174.00	15.00	104.00	29.00	23.00	3.00
	PER MIL	0.37	0.16	1.15	0.30	0.25	0.03
<i>buzz off</i>	FREQ	63.00	10.00	27.00	16.00	7.00	3.00
	PER MIL	0.14	0.10	0.30	0.17	0.08	0.03
<i>* pest</i>	FREQ	118.65					
	PER MIL	0.26					
<i>goalie</i>	FREQ	710.00	38.00	93.00	99.00	469.00	11.00
	PER MIL	1.53	0.40	1.03	1.04	5.11	0.12
<i>You bet</i>	FREQ	2,069.00	1,197.00	486.00	202.00	158.00	26.00
	PER MIL	4.46	12.53	5.37	2.11	1.72	0.29
<i>the States</i>	FREQ	2,613.00	512.00	587.00	544.00	340.00	630.00
	PER MIL	5.63	5.36	6.49	5.69	3.71	6.92
<i>howdy</i>	FREQ	430.00	124.00	199.00	58.00	42.00	7.00
	PER MIL	0.93	1.30	2.20	0.61	0.46	0.08
<i>crush</i>	FREQ	2,092.00	301.00	739.00	468.00	487.00	97.00
	PER MIL	4.51	3.15	8.17	4.90	5.31	1.07
<i>pretty</i>	FREQ	67,187.00	28,234.00	16,024.00	10,540.00	10,666.00	1,723.00
	PER MIL	144.69	295.44	177.20	110.30	116.29	18.92
<i>luv=love</i>	FREQ	3.00	1.00	2.00	0.00	0.00	0.00
	PER MIL	0.01	0.01	0.02	0.00	0.00	0.00
<i>cussing</i>	FREQ	37.00	3.00	20.00	6.00	8.00	0.00
	PER MIL	0.08	0.03	0.22	0.06	0.09	0.00
<i>mate</i>	FREQ	5,615.00	1,145.00	1,345.00	1,695.00	816.00	614.00
	PER MIL	12.09	11.98	14.87	17.74	8.90	6.74
<i>how come</i>	FREQ	2,806.00	937.00	1,364.00	291.00	169.00	45.00
	PER MIL	6.04	9.80	15.08	3.05	1.84	0.49
<i>laid back</i>	FREQ	288.00	77.00	74.00	47.00	79.00	11.00

	PER MIL	0.62	0.81	0.82	0.49	0.86	0.12
<i>a granny</i>	FREQ	1,830.00	140.00	1,070.00	380.00	201.00	39.00
	PER MIL	3.94	1.46	11.83	3.98	2.19	0.43
<i>a kid</i>	FREQ	36,058.00	8,820.00	13,727.00	6,641.00	6,308.00	562.00
	PER MIL	77.65	92.29	151.80	69.50	68.78	6.17
<i>gran</i>	FREQ	984.00	22.00	377.00	168.00	135.00	282.00
	PER MIL	2.12	0.23	4.17	1.76	1.47	3.10
<i>a stomach bug</i>	FREQ	12.00	4.00	2.00	6.00	0.00	0.00
	PER MIL	0.03	0.04	0.02	0.06	0.00	0.00
<i>crack a joke</i>	FREQ	64.00	8.00	13.00	17.00	25.00	1.00
	PER MIL	0.14	0.08	0.14	0.18	0.27	0.01
<i>knackered</i>	FREQ	18.00	1.00	14.00	3.00	0.00	0.00
	PER MIL	0.04	0.01	0.15	0.03	0.00	0.00
<i>gonna = going to</i>	FREQ	25,850.00	11,464.00	10,809.00	2,078.00	1,269.00	230.00
	PER MIL	55.67	119.96	119.53	21.75	13.84	2.53
<i>*a Yankee</i>	FREQ	292.57					
	PER MIL	0.63					
<i>a bit of</i>	FREQ	10,049.00	2,665.00	2,674.00	2,445.00	1,721.00	544.00
	PER MIL	21.64	27.89	29.57	25.59	18.76	5.97
<i>guy</i>	FREQ	77,105.00	28,734.00	20,751.00	13,910.00	12,958.00	752.00
	PER MIL	166.05	300.67	229.47	145.56	141.28	8.26
<i>the States</i>	FREQ	2,613.00	512.00	587.00	544.00	340.00	630.00
	PER MIL	5.63	5.36	6.49	5.69	3.71	6.92
<i>*Down Under</i>	FREQ	205.94					
	PER MIL	0.44					
<i>*mad</i>	FREQ	3,793.24					
	PER MIL	8.18					
<i>crazy</i>	FREQ	23,752.00	6,008.00	9,408.00	4,225.00	3,466.00	645.00
	PER MIL	51.15	62.87	104.04	44.21	37.79	7.08
<i>I guess</i>	FREQ	38,769.00	22,212.00	11,235.00	2,083.00	2,757.00	482.00
	PER MIL	83.49	232.43	124.24	21.80	30.06	5.29
<i>footy</i>	FREQ	5.00	0.00	2.00	2.00	0.00	1.00
	PER MIL	0.01	0.00	0.02	0.02	0.00	0.01
<i>*globetrotter</i>	FREQ	8.00					
	PER MIL	0.02					
<i>Aussie</i>	FREQ	349.00	29.00	53.00	135.00	124.00	8.00
	PER MIL	0.75	0.30	0.59	1.41	1.35	0.09
<i>stuff</i>	FREQ	47,152.00	18,078.00	10,972.00	9,128.00	7,423.00	1,551.00
	PER MIL	101.55	189.17	121.33	95.52	80.93	17.03
<i>*the box</i>	FREQ	0.00					
	PER MIL	0.00					
<i>an airhead</i>	FREQ	63.00	8.00	29.00	17.00	8.00	1.00
	PER MIL	0.14	0.08	0.32	0.18	0.09	0.01
<i>a pal</i>	FREQ	3,453.00	513.00	1,051.00	1,162.00	594.00	133.00
	PER MIL	7.44	5.37	11.62	12.16	6.48	1.46
<i>a kid sister</i>	FREQ	115.00	9.00	64.00	22.00	20.00	0.00

<i>*a jerk</i>	PER MIL	0.25	0.09	0.71	0.23	0.22	0.00
	FREQ	1,312.70					
	PER MIL	2.83					
<i>No way!</i>	FREQ	296.00	32.00	184.00	46.00	30.00	4.00
	PER MIL	0.64	0.33	2.03	0.48	0.33	0.04
<i>an ad</i>	FREQ	12,068.00	3,814.00	1,266.00	2,476.00	3,252.00	1,260.00
	PER MIL	25.99	39.91	14.00	25.91	35.46	13.84
<i>wacko</i>	FREQ	228.00	84.00	74.00	28.00	40.00	2.00
	PER MIL	0.49	0.88	0.82	0.29	0.44	0.02
<i>ad</i>	FREQ	12,068.00	3,814.00	1,266.00	2,476.00	3,252.00	1,260.00
	PER MIL	25.99	39.91	14.00	25.91	35.46	13.84
<i>advert</i>	FREQ	49.00	5.00	20.00	5.00	0.00	19.00
	PER MIL	0.11	0.05	0.22	0.05	0.00	0.21
<i>advertisement</i>	FREQ	1,669.00	232.00	232.00	379.00	330.00	496.00
	PER MIL	3.59	2.43	2.57	3.97	3.60	5.45
<i>acquire</i>	FREQ	6,176.00	445.00	429.00	1,227.00	1,274.00	2,801.00
	PER MIL	13.30	4.66	4.74	12.84	13.89	30.76
<i>porn</i>	FREQ	1,735.00	472.00	281.00	586.00	302.00	94.00
	PER MIL	3.74	4.94	3.11	6.13	3.29	1.03
<i>thus</i>	FREQ	62,764.00	1,308.00	3,281.00	9,678.00	3,986.00	44,511.00
	PER MIL	135.17	13.69	36.28	101.28	43.46	488.78
<i>a telly</i>	FREQ	233.00	20.00	156.00	23.00	32.00	2.00
	PER MIL	0.50	0.21	1.73	0.24	0.35	0.02
<i>*a box</i>	FREQ	0.00					
	PER MIL	0.00					
<i>on the go</i>	FREQ	425.00	78.00	33.00	233.00	66.00	15.00
	PER MIL	0.92	0.82	0.36	2.44	0.72	0.16
<i>vegetation</i>	FREQ	4,014.00	112.00	509.00	1,247.00	426.00	1,720.00
	PER MIL	8.64	1.17	5.63	13.05	4.64	18.89